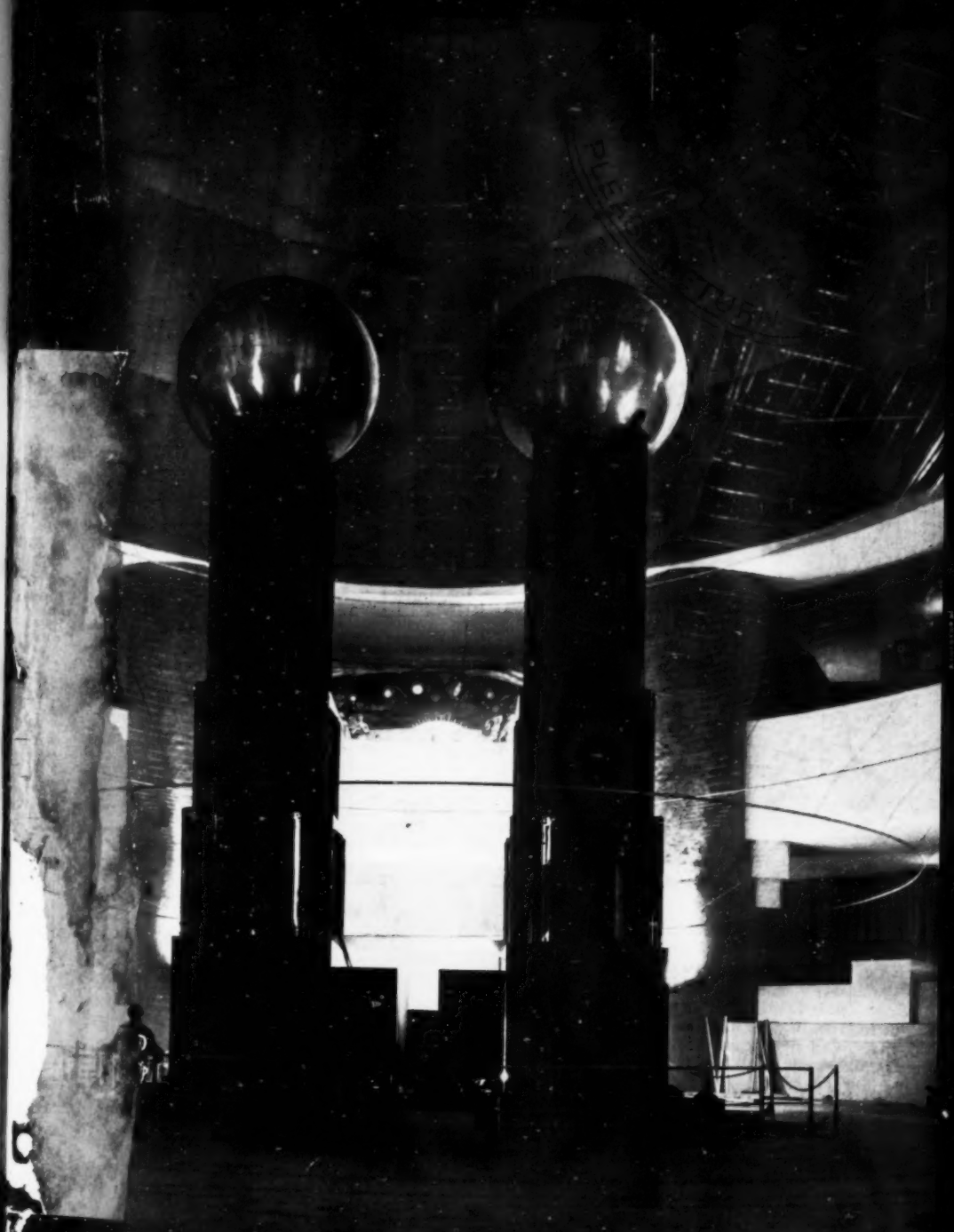
# JOURNAL OF APPLIED PHYSICS



JOURNAL OF APPLIED

# PHYSICS

(Formerly PHYSICS)

MAY, 1938 Vol. 9, No. 5

## *In This Issue*

*Page*

**"Palace of Discovery" at the Paris Exposition of 1937** JESSE W. M. DUMOND 289

**Reviews of Recent Research** 295

**Recent Progress in Supersonics** WILLIAM T. RICHARDS 298

**Study of Extreme Pressures and their Importance in the Investigation of Engineering Problems** THOMAS C. POULTER 307

**Application of the Lagrangian Equations to Electrical Circuits** D. A. WELLS 312

*Contributed Original Research*

**Contribution to the Problem of Arcback in High Voltage Discharge Devices** H. KLEMPERER 326

**Effect of High Energy Electron Random Motion Upon the Shape of the Magnetron Cut-Off Curve** E. G. LINDER 331

**Numerical Solution of LaPlace's Equation** G. H. SHORTLEY AND R. WELLER 334

*In Every Issue*

**Here and There** 321
**Innovations in Instruments** 323
**Some of the New Books** 324
**Contributors to this Issue** 325
**Current Literature of Physics** viii
**Advertising Index** xx

**BOARD OF EDITORS**

ELMER HUTCHISSON, *Editor*

*Associate Editors*

R. M. BOZORTH
D. W. BRONK
W. P. DAVEY
O. S. DUFFENDACK
SAUL DUSHMAN
A. STUART HUNTER
C. H. KUNSMAN
J. T. LITTLETON
L. L. NETTLETON
W. B. NOTTINGHAM
A. R. OLPIN
JOHN T. TATE

**Publication of the American Institute of Physics**

**ADMINISTRATIVE STAFF**

HENRY A. BARTON, *Director*

MADELINE M. MITCHELL, *Publications Manager*

WILLIAM C. COPP, *Advertising Manager*

**GOVERNING BOARD**

JOHN T. TATE, *Chairman*
KARL T. COMPTON
E. C. CRITTENDEN
WHEELER P. DAVEY
HOMER L. DODGE
HARVEY FLETCHER
ARTHUR C. HARDY
VERN O. KNUDSEN
MELVIN MOONEY
FREDERIC PALMER, JR.
R. L. PEEK, JR.
GEORGE B. PEGRAM
W. B. RAYTON
F. K. RICHTMYER
WALLACE WATERFALL

THE JOURNAL OF APPLIED PHYSICS, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as PHYSICS. It aims to be of service not only in physical laboratories, but also in the laboratories of industrial, chemical, geological, geophysical, meteorological, radio, and similar concerns. THE JOURNAL OF APPLIED PHYSICS publishes editorials, news of physicists and reviews, as well as technical papers of applied physics.

SUBSCRIPTION PRICE

| | U. S., POSS. AND CANADA | FOREIGN |
|---|---|---|
| To members of scientific societies | $5.00 | $5.70 |
| To all others | 7.00 | 7.70 |

**Manuscripts** should be submitted to Elmer Hutchisson, Editor, University of Pittsburgh, Pittsburgh, Pennsylvania.

**Changes of address,** proofs, and complaints of failure to receive the JOURNAL OF APPLIED PHYSICS should be addressed to the Publications Manager. New copies can be sent free only if notice of nondelivery is received within three months of date of issue.

**Subscriptions** and orders for back numbers should be addressed to Prince and Lemon Streets, Lancaster, Pa. or to the American Institute of Physics, 175 Fifth Avenue, New York, N. Y.

**Advertising rates** supplied on request. Orders, advertising copy and cuts should be sent to the American Institute of Physics.

Entered as second class matter at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1897. Accepted for mailing at the special rate of postage provided for in the Act of February 28, 1925, authorized May 2, 1932.

Copyright 1938, by the American Institute of Physics.

Journal of Applied

Physics

Volume 9 May, 1938 Number 5

# The "Palace of Discovery" at the Paris Exposition of 1937

By Jesse W. M. DuMond

*California Institute of Technology, Pasadena, California*

DURING the summer of 1937 it was my pleasure and privilege to visit on many occasions the "Palace of Discovery" of the Paris Exposition. This vast, splendid and instructive exhibit of scientific principles and progress was essentially the idea of that "grand old man" of French science, Jean Perrin. The details were admirably worked out with the wholehearted cooperation of many of the most important French scientific workers who gave freely of their time and energy.

The purposes of M. Perrin in organizing this scientific exhibition are so well expressed in his own words that I merely quote him here in my own translation.

"It should be made evident to all that the recent and prodigious progress of our civilization has had its source in *pure* and *disinterested* research directed toward the discovery of the unknown. Thus, no one should leave this exhibit without realizing that no part, I repeat absolutely no part, of the vast electrical industry of today would exist were it not for the discoveries of the electrical currents and their properties by the Italian, Volta; the Frenchman, Ampere; and the Englishman, Faraday.

"We wish at the same time to disseminate the idea that this country, that any country, has a primordial *and very practical* interest, however great its financial difficulties may seem, in giving to its first class research workers the sums of money, relatively insignificant as they are compared to a national budget, which they need for their research instead of economizing absurdly on the seed with which future harvests are sown.

"Finally pinning our hopes on the plain people, who in every country have shown in the most touching and moving manner their faith in the value of Science, we wish to stress among the adolescents who come with eager curiousity to see our experiments, the opportunities of vocational choice which may lead to the repetition of such miracles as the rise of Faraday from the position of book-

FIG. 1. Mounting of cathode ray to show sound waves.

binder-workman to that of the most eminent physicist of his country.

"The 'Palace of Discovery' built for this triple purpose was organized by *French* scientists. Those who visit it will see, however, that it is as *international* as is the science which it summarizes. Save for any forgetfulness or omissions which it shall be our constant effort to correct, we have tried to display the work of the great research men of all nations."

M. Perrin goes on from here to point out the solidarity of effort in world research which becomes even more apparent when we consider present living workers. For example, a discovery in nuclear chemistry may frequently be the result of the conscious collaboration of researchers in far distant lands all attracted at the same time toward the same enigma. He concludes by expressing the hope that a crusade in favor of scientific research such as the recent one in Belgium may spread to all countries in order that they may reap both the ideal intellectual and natural benefits that are sure to follow and he closes with the sentence, "Do you not see that then Peace will follow of necessity?"

The Science Exhibit was located in the "Grand Palais," the huge building which has in the past always been used for the annual exhibitions of French painting and sculpture. It was divided into sections of mathematics, physics, chemistry, astronomy and the biological sciences. It is impossible in a short article to do justice to all of these different sections. The excellent French journal *La Nature* in its issues of the first and fifteenth of September, 1937, has given fairly complete and accurate descriptions accompanied with many photographs and diagrams of the exhibits and experiments in the sections of physics and chemistry, respectively. I wish here to speak especially of the exhibits classified under physics. Taking for example the sections of mechanics and structure of matter, these were divided into sub-sections with the following names: Mechanics or Hall of Galileo (M. A. Marcellin), Crystallography (Mauguin and Wyart), States of Matter (J. Basset), Molecular Magnitudes (N. Choucroun), Fluorescence and Phosphorescence (F. Perrin, Rouault and Coustal), Negative Electrons and X-Rays (H. Hulubei, G. Y. Cauchois), Cosmic Rays (P. Auger and P. Ehrenfest), Radioactivity (M. Debierne), Atomic Synthesis (F. Joliot). I have purposely given after the title of each section the name of the French physicist who devoted time and energy to make it a success although these names were nowhere in evidence in the exhibit. It seems to me that it was this wholehearted and self-sacrificing devotion and cooperation of the French representatives of pure science and research which is the most instructive object lesson of the exhibit. Their work would of course have been sadly handicapped had they not been generously supported with funds and technical assistance.

One could, of course, go on to mention all the other divisions of physics, such as geometrical optics (Fabry, Bayle, Arnulf, Guerian, Penciolelli), physical optics and the velocity of light (Cotton, Andant), the Hall of Ampere and Faraday (Cotton, Tsai, Leprince-Ringuet, Poirson); oscillatory phenomena (Langevin, Lucas),

*THE photographs in this paper and the one used on the cover were supplied by the editors of the French periodical "La Nature," and we take this opportunity to express our appreciation for their kindness and cooperation.*—THE EDITOR

physics of the globe (Maurin, Salles) with their various subdivisions but space does not permit. The place was a veritable physicists' paradise—a show appealing at once to the most highly sophisticated savant and to the general public—there was a wealth for everybody. I shall mention therefore just a few exhibits and experiments taken at random.

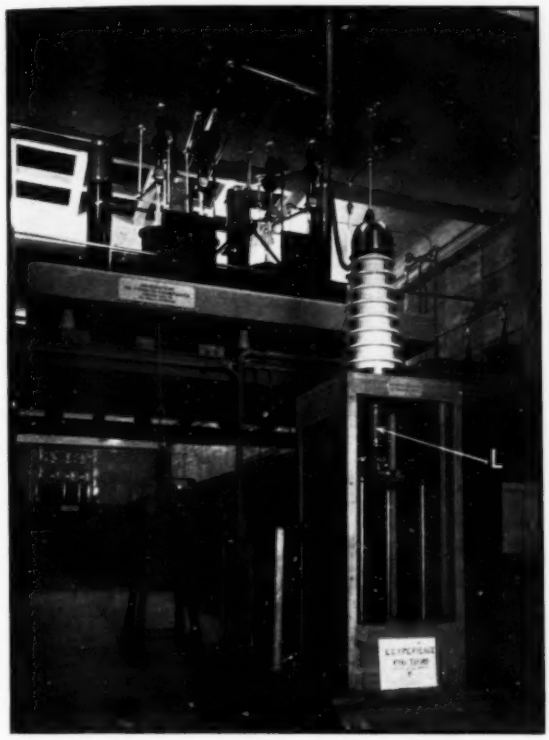



FIG. 2. General view of the "Hall of Negative Electrons" with the Lenard ray tube in the immediate foreground operated by the 250 kv 30 milliampere constant tension generator seen above. Fast electrons issue from the thin window in the bottom end of the tube *L* where they provoke luminescence in the air and in minerals which are placed below and which exhibit superb colors. Booths containing automatically operated exhibits can be seen lining either wall of this large room.

In the Hall of Galileo not far from the Foucault pendulum is a large rotating circular table turning night and day. Small plumb bobs and vessels of fluid placed at different radial distances from its axis of rotation showed that a citizen of this little rotating world would at each position form a quite different impression of the vertical and horizontal from that formed by the stationary observers in the room. The lesson was driven home with curious vitality by the fact

FIG. 3. Section of the large display of crystal models.

that *grass* and other plants planted in earth distributed all over the surface of the rotating disk had grown in the direction of the plumb bobs and normal to the water surfaces but inclined relative to the earth's vertical by greater angles the farther the plants and grass were from the center. The effort of growth of the plant was always exerted against what for it was the prevailing gravitational field. Thus was both centrifugal force and Einstein's principle of equivalence driven home.

I cannot forget the demonstrations of change of state, the critical point, the continuous passage from gaseous to liquid states, etc. so ably explained to the spectators. Nor can I pass over the elegant microprojections permitting anyone to see the Brownian movement, discontinuous thickness of thin films, etc., and to understand through clear explanations the evidence for the reality of the molecular structure of matter. The entrancing new fields of monomolecular layers or two dimensional fluids furnished some delightful experiments. Such layers apparently can behave as two-dimensional solids (as Devaux has shown) or as two-dimensional liquids or gases (Marcellin, Langmuir, Blodgett).

Under crystallography a bewildering multitude of models made to a uniform scale of 400 million times magnification showed how x-rays reveal the structure of materials and teach us most valuable practical information as to why, for example, diamond is hard and graphite is soft though they are made of exactly the same atoms. The value of this discovery *in pure science* for

the understanding of nature, even in biology or in the practical field of the invention of new materials with desired physical properties is a

FIG. 4. Wilson cloud expansion chamber to demonstrate ranges of alpha-rays. Built by Surrugue.

lesson whose value we think should sink deeply into the minds of those who hold the purse strings of potential funds for research.

The highly practical sections on electrolysis and electrophoresis as well as the section on photoluminescence can only be mentioned with warm praise for lack of space. However, the Hall of Negative Electrons was of such exciting interest to me, I must tell about it in some detail. Mlle. Cauchois and M. Hulubei are to be congratulated for the splendid results of their really very ambitious efforts. Nearly all the exhibits which lined the walls in booths on either side of a large room were *automatic* so that the visitor, by pushing appropriate buttons and levers and reading the accompanying printed explanations, could instruct himself as to the nature and laws of thermionic emission, the deflection of electron beams by electric and magnetic fields, the ways of determining the specific charge $e/m$, photoelectric emission, electric discharges in different gases at different pressures, x-ray generation, x-ray absorption, photographic effects of x-rays, ionization of gases by x-rays, and many other experiments in generous profusion. Especial mention should be made also of the contributions to this exhibit from the work of Manne Siegbahn and of the electron diffraction tube operating automatically which J. J. Trillat of Besançon set up. The latter was arranged so that by pushing a button anyone could see as often as he liked a beam of electrons diffracted by a crystal forming beautiful clear and regular "Laue" spots on a fluorescent screen. The important results which Cauchois and Hulubei have achieved in the development of the curved crystal focusing x-ray spectrograph which has recently permitted the discovery of element 87 by these workers were exhibited along with many examples and forms of this powerful new tool. As a climax there was in the center a large high voltage cathode-ray tube fitted with a Lenard window from which could be made to issue into the air fast electrons which falling on an assortment of minerals made the latter emit a gorgeous display of colors.

For those who read without running printed explanations and spectra displayed about the "Hall of Negative Electrons" made evident such highlights of the history of atomic physics during

FIG. 5. Trillat's permanently evacuated and automatically operated electron diffraction tube. The fluorescent screen at the left end of the tube showed the von Laue patterns of the electron waves diffracted by a crystalline lattice very clearly to all spectators who pressed the button.

the twentieth century as the Bohr atom model, the concept of atomic number of Moseley, the wave particle duality of electrons and of radiation based on the immortal experimental findings of Gray, Compton, Davisson and Germer, and G. P. Thomson together with its enunciation in Heisenberg's principle of indetermination.

One is torn between the limitations of space and the aversion for omitting to speak of even one of so many well conceived displays; the section on radioactivity and atomic synthesis, for example, with its Wilson cloud chambers in operation and with the original apparatus of the Curies; the section on cosmic rays with the automatically started cloud chamber set off by counters so skillfully realized by Auger and Ehrenfest; the projection views accompanied with automatic phonograph transcriptions explaining the epoch making discoveries and photographs taken with the Wilson chamber by

FIG. 6. High luminosity spectrographs and some models of organic chemicals from the chemistry section.

workers in this field from America, England, France and many other countries with a fine, generous impartiality. There was the huge electrostatic belt generator of *Van de Graaff* type which formed the "pièce de resistance," as one entered the great central rotunda of the exhibit which so impressed the French public that it was very amusingly used in political cartoons in the newspapers in which the debates in the Luxembourg over financial policy were represented as the crashing disruptive discharges of this machine.

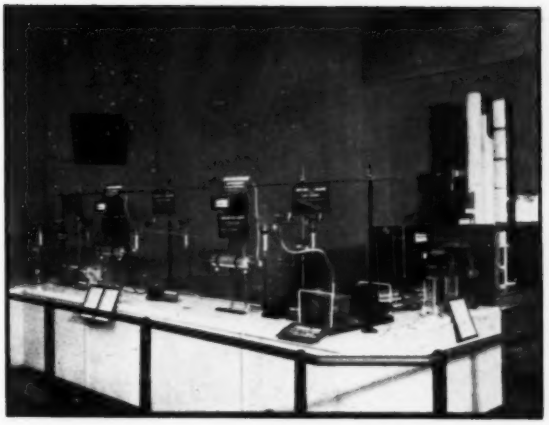


FIG. 7. The constituents of air including the rare gases were separated and exhibited by means of electric discharge tubes continually for every spectator to see in this apparatus. Section on chemistry.

By the time one had reached the section on optics one was prepared for anything and again one was not disappointed. There were beautiful and lucidly clear explanations of geometrical optics showing the rays going through lenses or reflected by mirrors or showing the formation of the mirage or of the rainbow; the experiment of the measuring of the velocity of light with mirrors turning on a jet of air and a beam only thirty meters long; every type of interference and diffraction experiment of which it seemed to me I had ever heard and many others all in perfect adjustment and clearly explained so as to make the wave properties of light self-evident, and a host of other beautiful experiments making clear the scattering of light, the colors of the sunset, the polarization and rotatory polarization of light, double refraction and its multitudinous applications, and so on almost it seemed without end.

The Ampere-Faraday hall contained a huge so-called "homopolar" generator which I believe was said to furnish 50,000 amperes in some experiments. These tremendous currents flowing through huge conductors were made to perform weird tricks at various places in the room and visitors were warned to leave their watches at a distance. Measurement of the magnetic fields with the "Cotton balance," a clear demonstration of the Zeeman effect, experiments on the magnetic birefringence of liquids, on falling bodies in a magnetic field and a very large

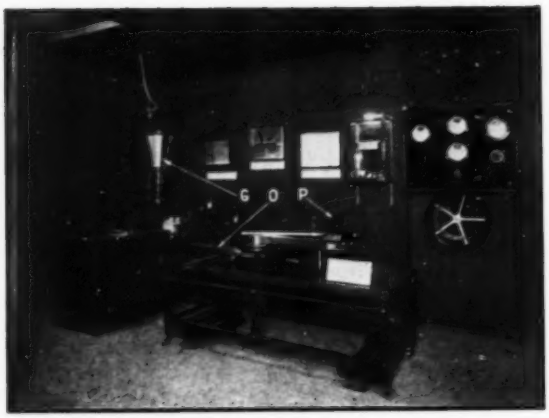



FIG. 8. Cauchois curved crystal spectrograph for demonstrating absorption and emission x-ray spectra. *G*, x-ray tube; *O*, optical bench; *P*, photographic plate holder. From the "Hall of Negative Electrons."

magnetic Wilson cloud chamber of de Broglie and Leprince-Ringuet were a few of the other exhibits which used the source of the immense electrical current just mentioned.

Of the sections on sound and ultrasonics and other vibratory phenomena we have not room to speak nor can we say but a word about the fine section on meteorology where there were striking experimental demonstrations of the cellular formation of thermoconvective vortices.

And we must not forget that my description says nothing of the subdivisions other than physics; the great display for instance of the progress in biology, medicine, surgery, heredity, microbiology, etc., etc., nor can I speak of the subdivisions of astronomy nor of mathematics with its museum of precise mathematical instruments or its little round treasure room whose frieze was decorated with the value of $\pi$ computed to heaven only knows how many decimals spiralling again and again around. And last but not least the subdivisions on chemistry and industrial chemistry so vital to our modern civilization must be passed undescribed.

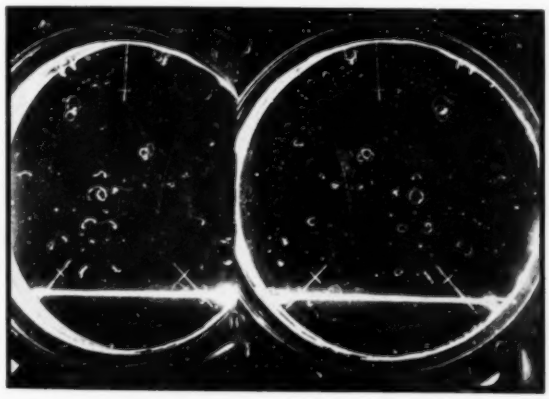
But I have said enough perhaps to point out the moral to this tale which many a reader may already have guessed. In this work-a-day world I think recently one can detect a misunderstanding which may have grown out of the too speedy and careless popularization of science—a misunderstanding of the nature, value, and role, as Perrin put it, "*of pure and disinterested research directed toward the discovery of the unknown.*" There is no doubting the meaning of his well chosen words. As our own Jewett has said "You can't have applied research until you have built up by pure research a body of knowledge to apply." Experience has shown (and an exhibit such as the French "Palace of Discovery" is a marvelous way of teaching this) that these frontiers of the known which must be ceaselessly pushed forward in order that all our technique of civilization may flourish and grow in its wake *can only be pushed forward by disinterested research. The immediate practical so-called utilitarian goal* must therefore not be our exclusive aim or we risk nipping in the bud the Newtons and the Faradays and the Maxwells and the Diracs and the Heisenbergs. It is the lesson of their part in the world picture that such a science exhibition should aim to teach. Whether we believe that government (as in France) or private initiative (as in America) is the best agency to encourage pure research is certainly a question which I believe is still open for debate but it is certain to me in either case that the value of "pure and disinterested research directed toward the discovery of the unknown" must be driven home to the public unless civilization is to die a slow death from malnutrition in these United States. The physicists of America have two opportunities to do their part in this cause in the near future—one is the exhibition in San Francisco, the other the exhibition in New York. Let us hope they will see the need and rise to the occasion.

FIG. 9. Reflection, refraction, and grating diffractions of short radio waves.

# Reviews of Recent Research

## Study of Gamma-Radiation from Boron Bombarded by Protons

THE knowledge of the properties of the excited states of atoms and molecules secured by experiment forms a large part of the basis of the present theory of atomic and molecular structure. It is now believed that experimental data on the properties of nuclear states of excitation will aid in developing and testing an adequate picture of nuclear structure. One method of determining the properties of excited nuclear states is by studying the intensity and energy of the gamma-radiation emitted in transitions to lower lying states. Those transmutations involving the capture of fast protons by light nuclei are especially useful in this connection as they result in the production of excited states of only one nucleus or at most two if the bombarded element has two isotopes. In various laboratories, studies of the gamma-radiation from transmutations have been attempted by securing cloud chamber photographs of secondaries (Compton recoil electrons and electron-positron pairs) produced in thin sheets of material such as lead, carbon, or aluminum by the radiation in question.



An electron-positron pair produced in the gas of a cloud chamber by a gamma-ray of 16.1 Mev energy produced in turn by bombarding boron with fast protons. The electron member of the pair bending toward the intersection at the two stereoscopic views of the cloud chamber has 9.5 Mev kinetic energy while the positron has 5.6 Mev kinetic energy. The production of the pair required 1.0 Mev of the gamma-ray's energy. The curvature is due to a magnetic field of 2500 gauss.

Recently* W. A. Fowler, E. R. Gaerttner, and C. C. Lauritsen of the W. K. Kellogg Radiation Laboratory of the California Institute of Technology have found in this way that when boron is bombarded with protons, gamma-rays of energy $16.6 \pm 0.6$ Mev, $11.8 \pm 0.5$ Mev, and $4.3 \pm 0.3$ Mev with relative intensity $1/7 : 1 : 1$ are emitted. The nucleus formed in this case is an excited state of carbon which can decay to normal carbon directly or in a double transition through a state at 4.3 Mev above normal carbon. This state has been found to be produced in numerous other transmutations.

★

## Penetrating Power of High Velocity Electrons Studied

ELECTRONS fast enough to travel 130 meters in air can now be produced in abundance by means of artificial nuclear disintegration processes. These serve to fill in the gap between the extreme energies found in cosmic rays (which easily traverse the entire atmosphere of the earth) and the much lower energies of the naturally occurring radioactive substances. Experimenters can therefore make measurements of the phenomena associated with the penetration of electrons through matter over the entire spectrum of energies.

During the past two years detailed measurements have been carried out at the University of Michigan to determine the ability of electrons of energies from 0.5 to 17 million electron volts to penetrate sheets of lead and carbon. This energy range is especially interesting because it includes a transition from one type of absorption to another. Theory predicts the following: At the lowest energies the loss of energy suffered

* Phys. Rev. **53**, 628 (1938).

by an electron is due almost entirely to the production of ions along its path or to collisions with other electrons. At very high energies the latter is entirely overshadowed by losses due to the production of x-rays or high energy quanta. When lead is used as the absorbing material the transition between these two kinds of absorption falls within the energy range studied: for example, at 3 Mev the loss is due 75 percent to ionization, at 10 Mev it is 50–50, and at 30 Mev it is due 75 percent to radiation. The values for the losses found in the Michigan experiments, the latest of which are reported by A. J. Ruhlig and H. R. Crane in the April 15th issue of the *Physical Review*, are higher than the theoretical predictions over the entire range of energy studied. They run from 3 to 4 times the theoretical loss at 0.5 Mev to about 1.5 times the theoretical loss at 15 Mev, and from the trend of the values it appears that as the energy is increased the agreement between experiment and theory will become better. One of the difficulties in the interpretation of the experimental results is that the electrons pursue crooked paths inside the absorbing material, making it hard to estimate the true distance traveled. This is particularly true for low energy electrons, and may very well account for at least some of the apparent excess in the energy loss.

Cloud chamber photograph showing the change in curvature of an electron path before and after passing through a sheet of lead. Electrons enter the chamber from the lower edge of the picture.

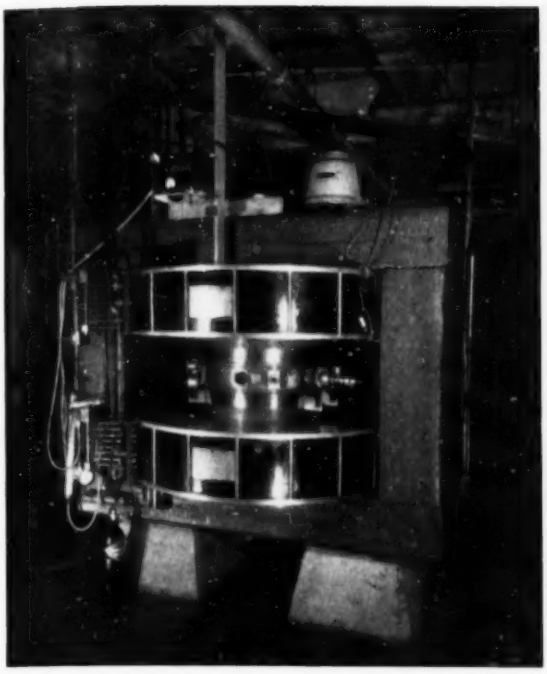
The cloud chamber picture shown in the figure will give an idea as to the method of experiment. Electrons, whose paths are curved by a magnetic field, are seen entering the chamber from the lower edge of the picture. Some of them pass through the sheet of lead, which appears as a horizontal line across the center. The curvatures of the paths before and after passing through the lead tell how much energy each electron had to begin with and how much it lost in traversing the lead. Many times electrons "bounce" backward from the lead. This is due to one or more large deflections caused by close approaches of the electron to lead nuclei. The two electrons entering 2nd and 3rd from the left in the figure are deflected about 180 degrees with very small loss in energy. The electrons shown in the figure have energies of the order of 4 Mev.

★

## Proton Induced Radioactivities

WHILE most of the cyclotrons now in operation in this country have confined their attention to the nuclear disintegrations produced by high speed deuterons and alpha-particles, the one at the University of Rochester* has been devoted to the investigation of reactions produced by high energy protons. Using a proton beam whose energy is 4 Mev (which has more recently been increased to 5.5 Mev) a large number of radioactive isotopes have been produced. The particular feature of this work is that a previously unknown type of reaction has been discovered in which a proton entering a nucleus ejects from it a neutron. In almost all cases the resulting nucleus is radioactive, in some cases emitting positrons and in other cases, electrons. In this paper results are reported on the radioactivities observed in proton bom-

* Described by L. A. DuBridge, *et al.* Phys. Rev. **53**, 447–454 (1938).



The cyclotron of the University of Rochester with a 27-inch chamber. It is rated to produce 7 Mev protons.

bardment of the following elements: N, O, Cr, Co, Zn, Se, Mo, Pd, Cd, In. Particularly strong radioactivities were observed in targets of oxygen and selenium and the energy relations involved in these reactions have been studied in some detail.

The positron emitting isotopes formed by the proton-neutron reaction are of particular interest. In this case after decay the radioactive isotope returns to the initial isotope. The mass relations are very simple and lead to the prediction that the minimum proton energy at which the reaction may take place is equal to 1.8 Mev plus the maximum energy of the emitted positrons. This relation has been experimentally verified.

★

## Electrostatic Generator Operating Under High Air Pressure

A RECENT paper in the *Physical Review*[1] describes improvements in the high voltage generator developed at the University of Wisconsin for experiments in nuclear physics. The generator is of the electrostatic type, and operates in a horizontal cylindrical tank 20 feet long and $5\frac{1}{2}$ feet in diameter, under an air pressure of 100 lb./in.$^2$, this pressure giving a maximum voltage four times as high as that obtained with air at atmospheric pressure. The addition of a few kilograms of $CCl_4$ or $CCl_2F_2$ vapor to the air in the tank has been found to increase the maximum usable voltage from 2.1 to 2.4 million volts; so that the upper limit to the voltage is now set by a breakdown along the accelerating tube, the charging belts, or the supporting insulators rather than by radial sparking across the 18-inch air gap between the high potential electrode and the wall of the tank. Generator voltage can be easily set at any desired value up to the maximum and is held at a given value to within less than 1 percent by an automatic voltage stabilizer.

Two balloon fabric charging belts running along the axis of the tank from one end feed charge to the high potential electrode which is located near the center of the tank. Extending from this electrode along the axis of the tank to the opposite end is the evacuated accelerating tube, which consists of a number of corrugated porcelain cylinders separated by metal electrodes. Protons from the high potential electrode travel down the accelerating tube, and pass out through the evacuated tube at the end of the tank, where they are at present being used for proton-proton scattering experiments, and for measurements on nuclear transmutations caused by proton bombardment.

The Wisconsin electrostatic generator.

[1] D. B. Parkinson, R. G. Herb, E. J. Bernet and J. L. McKibben, Phys. Rev. **53**, 642 (1938).

# Recent Progress in Supersonics*

By WILLIAM T. RICHARDS

*Rockefeller Institute for Medical Research*

*New York, New York*

I MUST begin with an apology. In going over the development of supersonics during the past five years I found that it is quite impossible, in a short address, to give credit where credit is due in all cases. What I am about to say, therefore, will be an attempt to trace the development of several main lines of investigation, and much interesting work which is not intimately associated with these main lines cannot be mentioned. I hope that any of you who consider that I have sinned either by omission or by commission will make your views known to me as loudly and as emphatically as possible. In an expansion of this address, which will later appear in the *Reviews of Modern Physics*, I shall give a more scholarly exposition. I shall speak first of the propagation of sound in gases, liquids and solids. Later I shall devote a few paragraphs to the chemical and physical effects of sound of high amplitude.

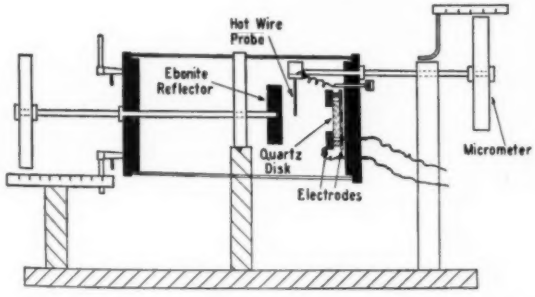
The theory of the propagation of sound in dispersive gases has been very fully treated by numerous writers. In general they follow either the statistical or the thermodynamic discipline. In general also they consider the dispersion due to the failure of a part of the heat capacity to follow the adiabatic cycle of the sound wave. To illustrate this as specifically as possible let us consider chlorine, a gas which is partly excited vibrationally at room temperature and which has, in consequence, a heat capacity at constant volume greater than $5/2$ of $R$ per mole. If the velocity of sound is measured at low frequencies, the rate at which the small pressure impulses of the sound waves succeed each other is less than the rate at which translational energy is converted into vibrational energy, and the effective heat capacity is the equilibrium or "zero frequency" heat capacity. In the "zero frequency" range the gas is dispersive only because of second order effects such as viscosity, heat conduction, or finite amplitude of the waves. When the frequency has been raised to within one order of magnitude of the rate of conversion of translational into vibrational energy, dispersion from this cause becomes manifest. The reason for the dispersion is the failure of some of the molecules of the gas to alter their vibrational energy to conform to the rapidly changing local temperature. This results in a diminution of the effective heat capacity. The pressure-temperature cycle of the sound wave remains adiabatic, but is no longer at maximum entropy. A phase angle has appeared between the propagation of the pressure and the density waves, and the tangent of this angle gives, of course, the absorption coefficient. At circular frequencies one order of magnitude or more above the rate of conversion of translational into vibrational energy, no appreciable effect of the vibrational energy remains. The velocity of sound has reached a new and higher value, which remains constant even at very high frequencies, and corresponds to a heat capacity of $5/2\ R$ for chlorine. Both the thermodynamic and the statistical treatment agree in describing this dispersion by the methods of modern chemical reaction kinetics, although this is not necessarily inherent in the thermodynamic viewpoint. Hence they arrive at essentially similar conclusions, which are currently termed "the collision theory of acoustical dispersion."

Bourgin is responsible almost single-handed for the development of the statistical theory. In 1928 he wrote the only complete theory of dis-

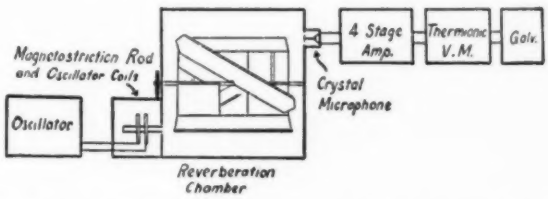
* An Address to the Acoustical Society of America at Ann Arbor, November 30, 1937. Read by the Society's president for the author.

persion for waves of infinitesimal amplitude that I have seen.[1] He considered viscosity, quantization of the collision process, radiation from excited molecules, and dispersive gaseous mixtures. He followed this by expanding his treatment of gaseous mixtures, and, in 1932, wrote the first theory of the three-state gas.[2] This is a gas having translational energy and two sets of internal energy states which are characterized by different rates of adjustment. In chlorine the two states would be the vibrational and the rotational energy. Finally, he has recently compared the statistical and the thermodynamic methods.[3] Considering the brilliance of this achievement Bourgin should be honored first among those writing on dispersion, yet not one experimenter has, so far as I am aware, taken advantage of this work. I think that this is only because the statistical method involves reasoning of considerably greater mathematical difficulty than the thermodynamic. We experimenters are simple folk, easily awed by a row of summation signs, and we have chosen the simplest language we could find. Professor Bourgin has had to wait ten years for the recognition which is his due. He may have to wait another ten. But since a statistical method is the most logical way of describing a statistical process, he must ultimately triumph.

The thermodynamic method was first published in articulate form by Einstein in a paper of great beauty and simplicity dealing with dispersion in a dissociating gas.[4] Kneser, ten years later, applied the same method to a pressure-density phase lag in a nondissociating gas.[5] Kneser made a serious error which Rutgers, leaning even more heavily on the exposition of Einstein, corrected.[6] I found it necessary to extend Einstein's treatment to a three-state gas in order to prove that it was the dissociation reaction and not the heat capacity which caused dispersion in nitrogen tetroxide.[7] Luck extended Einstein's treatment to describe absorption in a real dissociating gas,[8] and I followed his lead by correcting for van der Waals forces in nondissociating gases and gaseous mixtures.[9] I also worked out the behavior of four- and five-state gases, but this, judging by the profound apathy with which the publication[10] was greeted, was carrying matters too far. The work of Landau and Teller[11] makes it appear that such monstrosities will not be necessary. The method of Einstein—and as a reviewer I must insist that it be called the method of Einstein and not the Kneser-Rutgers method—has yet to be corrected for viscosity and heat conduction.




Richardson's hot-wire probe apparatus for measuring acoustic absorption in gases.

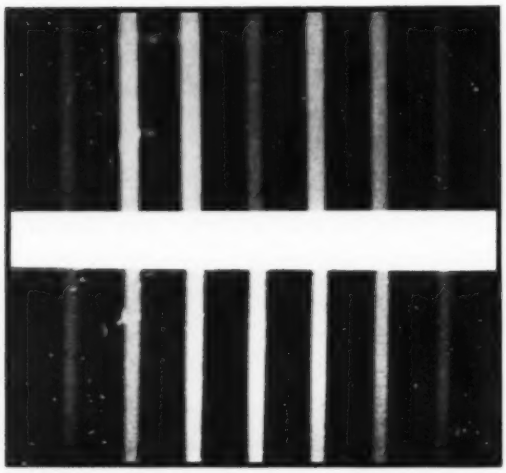



Knudsen's arrangement for measuring the absorption of sound in a gas by the intensity method.

The very stimulating contribution of Herzfeld and Rice,[12] which was prior to all the work mentioned above save only that of Einstein, does not fall readily into either of these two classifications, although it also postulates a collision mechanism.

It should be apparent from what I have said that a believer in the collision theory may find appropriate expressions to describe his measurements in either Bourgin's or Einstein's language. In both sufficiently explicit expressions have been written to enable him to identify the cause of any dispersion or absorption he may find, provided only that his experimental technique is equal to the occasion.

Fortunately, perhaps, for the vitality of the subject, some investigators, who have obtained results which they believe inconsistent with the collision theory, proceed to deny its validity. I think especially of Richardson[13] and of Penman[14] in England. I should be entirely in sympathy with this enterprise if I had not personally encountered far more startling irregularities which



The diffraction of light by sound waves. The upper figure shows the intensity distribution with progressive waves and the lower with standing waves. (After R. Bär.)

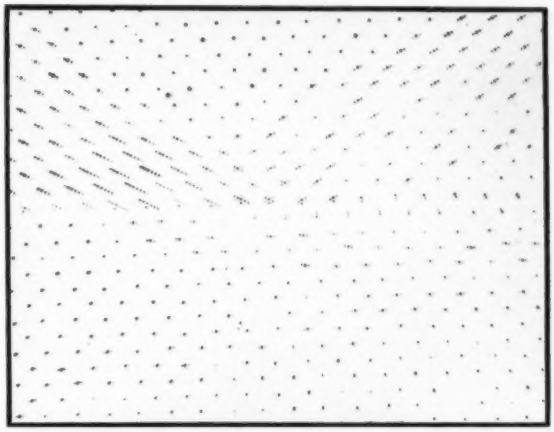
later proved to be apparatus effects. I shall mention one or two of these in a few minutes. A collective work entitled *Aberrations of Sonic Interferometers* would have a useful place in acoustical literature. In view of this I shall here take the high-handed stand that measurements which do not agree with the collision theory are experimentally unsound. If this stand can be refuted, my interest in the subject will be greatly refreshed.

Before leaving the theoretical aspect of the propagation of sound in gases I should like to call attention to a tendency on the part of experimenters to overidealize the situation. For example, expressions developed for ideal gases are often cheerfully used over wide temperature and pressure ranges. The most common of these abuses is the so-called "reduction to zero degrees centigrade" by treating the velocity as if it were a linear function of the square root of the absolute temperature. For carbon dioxide this is not even a close approximation. Indeed, "zero frequency" velocity measurements appear to offer an accurate means of evaluating the second virial coefficient. Again, there is the tendency to regard the waves generated by a piezoelectric or magnetostrictive oscillator as plane, purely sinosoidal, and of infinitesimal amplitude. They certainly have none of these characteristics to any high degree of approximation. Fay, in a valuable and little-read paper,[15] has shown that, for waves of even very moderate amplitude, the attenuation due to the conditions of energy transfer may be twenty times that due to viscous loss. This attenuation is due solely to the fact that the sound wave is a sound wave, and that the condensed portions are traveling faster than the rarefied portions in proportionality to the amount of matter in each. Pielemeier has concerned himself with the experimental aspects of this problem.[16] Further, we tend to neglect the effect of viscosity and heat conduction. This effect is small except in tubes having a diameter of the order of magnitude of the wave-length, but we should not forget that it is present. In cases like oxygen at room temperature, where the low and high frequency velocities differ by only a few tenths of a percent, it may be necessary to reckon with it. Finally, the possibility of radiation from a dipole as a means of energy loss must not be overlooked. In carbon dioxide the radiation must be slow, and has been shown experimentally not to affect the acoustical measurements. In other gases, however, which have large permanent dipoles and dispersive regions in low frequency ranges, it could well appear.

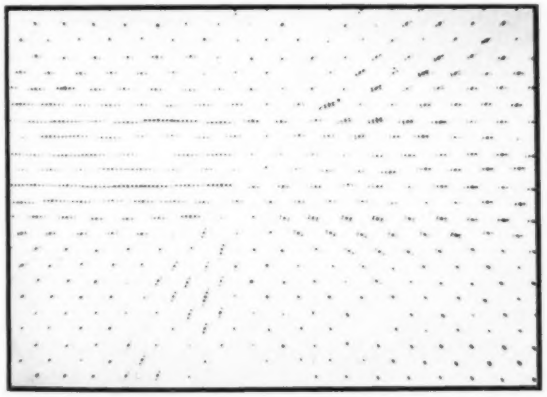
There is little to be said concerning the theory of sound propagation in liquids and solids. In liquids, the frequency of molecular collisions, if one can use such a term in a semi-rigid lattice structure, is so great that there is little chance of detecting dispersion due to heat capacity in frequency regions at present available for experiment. Moreover, if dispersion were found, its interpretation would present grave difficulties since the kinetic theory of liquids, though definitely conceived by the efforts of Fowler, Eyring and Bernal, is still in the foetal stage and cannot be expected to give assistance to other branches of the science. It appears that, conversely, measurements on the scattering of sound in liquids will lead to a substantiation of some of these theories, as Lucas[17] has recently suggested. Dissociating liquids, such as acetic acid, might be expected to show dispersion owing to the chemical reaction involved. Also there should be dispersion, closely analogous to the anomalous dispersion of radio waves, during the liquid-glass "transition." Similarly, in some crystals, where free rotation of certain groups begins below the melting-point, measurements of the dispersion of

sound should provide interesting information concerning the rate of energy transfer to these groups. I am not aware, however, that any of these phenomena have been reported.

By far the most popular instrument for measuring the velocity and absorption of sound is the sonic interferometer introduced by Pierce.[18] It has been modified for liquids and solids, and has been used to frequencies of half a megacycle with magnostrictive and twenty megacycles with piezoelectric oscillators. Nearly a hundred papers have appeared dealing with various aspects of this method alone. As evidence has accumulated it has become increasingly clear that it is not sufficient to set up an oscillator, calibrate it, and measure the position of points of maximum electrical reaction to secure numbers representing the velocity and absorption of sound which have absolute numerical significance. The wave pattern given off by the oscillator is not a pure plane sinosoid, since various parts of the surface are not vibrating in phase with each other. Grossmann has shown that, in consequence of this, a plane reflector will give maximum reaction on the oscillator at a position which is not exactly an even number of half-wave-lengths from the surface of the oscillator, but is a compromise with the various components of the wave pattern.[19] In addition, all parts of the oscillator are sending out waves, and, if measurements are made in a confined space, wall reflections occur. When all this is considered it is clear that very complicated reaction-distance patterns may be found, and that they must be interpreted only with the greatest caution. Observers are often deceived by a reproducibility of the order of 0.001 percent into a belief that the numerical accuracy is equally great. Another difficulty which is too little appreciated is the effect of varying pressure on the frequency of an oscillator. It does not appear reasonable that a change from 250 mm to 500 mm pressure of acetic acid vapor on a massive nickel tube a foot long should change its frequency by 300 cycles, yet I have observed this phenomenon repeatedly and reproducibly. Cases are on record where no greater increases of pressure than this have caused a high frequency crystal wholly to stop oscillating. No doubt the effect is a multiple one, involving not only the direct damping of the oscillator, but also adsorption phenomena of a chemical nature which alter the relation of the crystal and its mount. These difficulties notwithstanding, the sonic interferometer is an instrument of high precision. It is capable of giving absolute velocities to better than 0.1 percent, and absorption coefficients, owing especially to the examination of the reflection coefficients involved by Hubbard,[20] to about 10 percent.

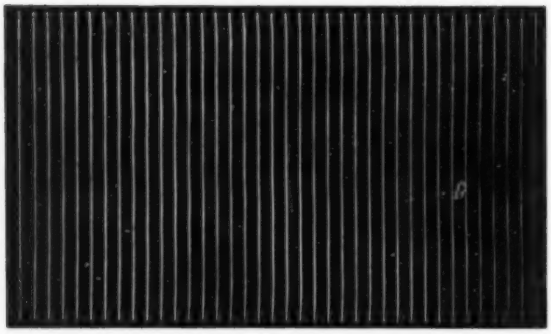


Diffraction and reflection of a supersonic wave in the surface between *p*-xylol (at the top) and water (at the bottom). Rays are incident from the northwest and are reflected to the northeast. Waves parallel to and just below the interface are going to the right. (After R. Bär.)



Diffraction of a supersonic wave by a wire grating. The piezoelectric source is on the left. The first-order diffraction spectrum is in a southeast direction, while the second-order spectrum is in a southwest direction. (After R. Bär.)

There are three welcome newcomers among the instruments suitable for measuring the velocity and absorption of sound in various media. One of these, which was devised by Debye and Sears,[21] employs the diffraction of light by the



Visibility of stationary ultrasonic waves in *m*-xylene. The frequency is 5 megacycles. This picture is by E. Hiedemann and Ch. Bachem.

alternate condensation and rarefaction of matter in the sound wave. Since standing waves are not required, the method is very simple, and it has been widely employed, especially by investigators interested in liquids. It is steadily undergoing improvement from a technical standpoint. The theory of the diffraction of light by sound waves has, in addition, attracted wide attention among mathematical physicists, the interpretations of Brillouin[22] and of Raman and Nath[23] being favored by most writers. The accuracy of this method appears to be almost as great as that obtainable with the Hubbard and Loomis[24] interferometer. It has the additional advantage of requiring small volumes of material; Bär has measured the velocity with only 1 cc of $D_2O$.[25]

A second new method is that using a movable hot wire probe which indicates regions of maximum and minimum motion by the temperature of the wire. This was devised by Richardson[13] and his co-workers, who recommend it especially for absorption measurements in gases. A thermopyle has also been used in this way. A systematic comparison of the performance of this instrument with the interferometer would be valuable. Since it has not been made, and since the hot wire probe has been used by few investigators, it is difficult to estimate its accuracy. A paper which has recently appeared in this journal describing an investigation of the nodal points in wind instruments by the hot wire method shows that it may be valuable in several branches of acoustics.

The most recent method depends upon the direct visibility of ultrasonic waves of even very moderate amplitude. With intensities as low as 2 watts per square centimeter standing waves of sound are visible to the eye in transformer oil. With the "Schlieren-Methode" they may be photographed at even smaller intensities in liquids, glasses, transparent crystals, and even in gases.[26] Such photographs may be used to determine the velocity of sound in conjunction with a microphotometer and a knowledge of the dimensions of the apparatus, gained either by direct measurement, or by comparison with a similar photograph of a substance of known properties.

The formation of a fog at liquid-air interface by intense high frequency sound waves. (Photograph by Wood and Loomis.)

The velocity of sound may in this way be measured in a hundredth of a second or less, which should be invaluable for a substance like formic acid vapor, which undergoes rapid thermal decomposition at temperatures where its vapor pressure is great enough to permit acoustical measurements. Bergmann, Heidemann, and others have recently developed interesting stroboscopic variations of this procedure. The important feature of these methods is that they sidestep any difficulties due to irregularities of the wave form.

The reverberation method, used with such brilliant success by Knudsen[27] for determining

absorption coefficients at audible frequencies has been extended into the ultra-audible range.[28] It would be aesthetically pleasing if a method could be found to measure the group velocity, since this would give the dispersion directly.

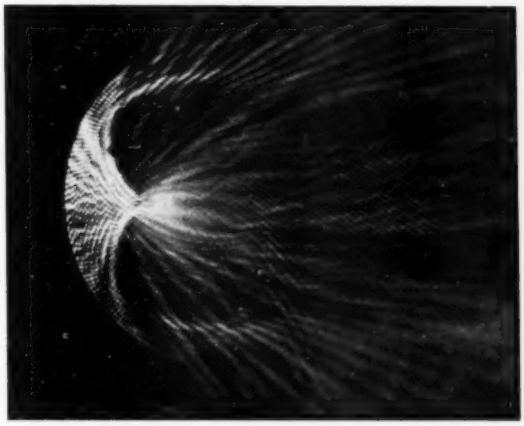
Turning now to the results of measurements, we must first consider the information to be obtained from the so-called "zero frequency" or equilibrium velocities. Gases and vapors, if a suitable compensation is made for the van der Waals forces, yield in this region data which should be a valuable confirmation of the analysis of band spectra. For carbon dioxide and carbon disulfide such confirmation is trivial, but for gases like ethylene, where the analysis is difficult, accurate measurements of the velocity of sound at low frequencies should be really helpful in view of the difficulty of determining the heat capacity by other methods. The spectral analysis of carbon dioxide, sulfur dioxide, chlorine, oxygen, and several other gases, has been checked in this way. In liquids, since the heat capacity cannot at present be evaluated theoretically, the significance of the velocity of sound is necessarily more superficial, and in fact appears to be valuable only for the determination of the adiabatic compressibility coefficient in nonelectrolytes. In electrolytes, since accelerated charges are involved, information concerning the masses of ions and their valences may also be obtained. Parthasarathy has been very industrious in measurements on a large number of organic nonelectrolytes in the hope of establishing relations between the velocity of sound and molecular constitution, but has not had conspicuous success.[29] The data on solids are meager.

By far the most attention has been focused on the dispersion and absorption of sound in polyatomic gases, since these data throw new and valuable light on the transfer of energy by impact. Too much has been published on this subject to permit its detailed discussion in the time at my disposal. There are, however, certain conclusions which result from the work of Knudsen, Kneser, Eucken, Pielemeier, Hubbard and their associates, and of Mr. Reid in my laboratory. I shall have to state these conclusions somewhat too categorically, but it is my belief that they are essentially correct.

(1) Dispersion and absorption of sound are to be expected in pressure, temperature, and frequency ranges at present accessible to experiment in any polyatomic gas in which an appreciable amount of vibrational energy is excited. The reason for this is primarily the inevitable inefficiency of the conversion of translational energy into vibrational energy by collision. In the vast majority of cases the effect of a collision will be merely to knock the molecules apart without bringing about any conversion at all. For the collision must convert one quantum of energy into vibration, not merely to perturb one of the two molecules sufficiently to set its atoms vibrating very slightly. Furthermore, a steric factor must enter, since a diatomic molecule must be hit on its end to produce vibration. These are crude statements, but I am told that a more detailed study of the collision process produces a very similar picture.

(2) The effect of increasing pressure is in accordance with the indications of the kinetic theory of heat. An increase of pressure causes more collisions per unit time, and consequently pushes the dispersive region to higher acoustical frequencies.

(3) The effect of increasing temperature is to increase exponentially the rate of conversion of translational energy into vibrational energy. The proportionality is not, however, that given by the Boltzmann factor for the particular energy state involved, but indicates that a larger energy than $h\nu$ is required. The energy in excess of $h\nu$ is



Pattern produced by the reflection of ultrasonic rays by concave mirror. After E. Hiedemann and Ch. Bachem.

roughly analogous to the activation energy of a chemical reaction, but it has been shown thoretically to depend on temperature, so the analogy cannot be taken seriously. Landau and Teller show that only molecules of high relative translational energy can effect conversion. Due to the phenomenon of "trace catalysis," which I am about to discuss, these temperature coefficients are very difficult to determine with accuracy. For this reason they are different, at present, in Germany from what they are in this country. I do not think that this should cause either surprise or distress.

(4) Owing to the pioneer investigations of Knudsen, which have been amply confirmed by others, it appears that very small traces of foreign gases may have enormous effect on the efficiency of the collision process, and consequently on the behavior of a gas in the dispersive region. The effect of the addition of these traces is always to shift the dispersive region to higher frequencies. Any other result would necessitate a revision of the collision theory. Kneser and Knudsen[30] interpret this result, I think correctly, as meaning that a collision of a molecule of the dispersive gas with a molecules of the "trace catalyst" is more effective in exciting the vibration of the dispersive gas than is a collision with one of its own species. The effect of changing concentration is not uniform, sometimes indicating that bimolecular collisions are involved, and sometimes trimolecular, and so on. Eucken[31] prefers to call this chemical excitation as opposed to physical excitation by impact with nonreactive gases such as He. Now phenomena like this have been familiar to chemists for a long time, and have distressed us greatly. If we are to think about it at all we must obviously go deeper than considerations of mere impact, since the average kinetic energy of all the molecules of the gas is the same. It is fairly reasonable to turn to rotational energy, and examine the molecules with small moments of inertia and consequently large rotational quanta. Chief among these is hydrogen, which appears to be an effective catalyst, and deuterium, which is only about one-eighth as effective as hydrogen with ethylene. Water and ammonia, and perhaps even methane, appear promising candidates, and do not disappoint us in experiments, although their interrelationships are not as simple as we could wish. But when Knudsen finds that ethyl alcohol is more effective with oxygen than any of a large number of trace catalysts he has tried, it becomes necessary to admit that rotational energy cannot explain, everything. For if, as Kistiakowsky and others have recently shown, there is no free rotation about the carbon—carbon bond in ethane, it is scarcely conceivable that there should be free rotation in ethanol, and the moment of inertia of the molecule, must be, therefore, very great. It seems to me, then, that we must turn in desperation to some idea of excitation by resonance between the sum of a set of frequencies in ethanol and the vibrational frequency of oxygen, for it is only in this way that I can envisage the adding together of several small quanta to make a large one. The torsional quanta of ethanol must be considerable, and these may participate in the exchange. During the collision the energy in a number of degrees of freedom in ethanol must flow over to the oxygen atom. A statement like this makes me uncomfortable, but its equivalent has been in use for many years in chemical reaction kinetics.

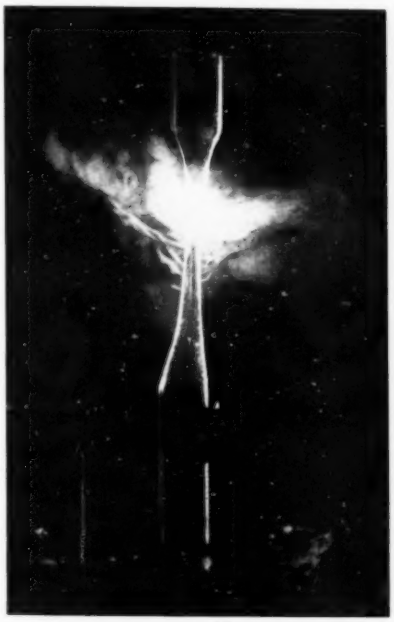
(5) No information has been obtained concerning the rates of adjustment of the rotational energy or of the van der Waals forces, since both appear to follow the acoustical cycle to the highest frequencies yet employed.

(6) Chemical reaction rates may be successfully determined by the acoustical method, but it seems that only gases with low heat capacities are suitable for investigation at present. Indeed the dissociation of nitrogen tetroxide is the only success to date, and this is a gas which is exceptionally easy to purify and manipulate, and has a low heat capacity as well.

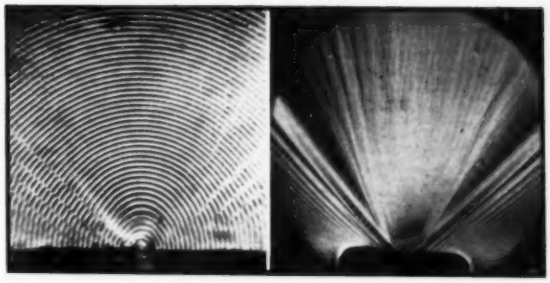
I shall close with a brief survey of the effects of intense sonic and supersonic radiation. This branch of the subject, which originated with the work of Wood and Loomis,[32] is in a somewhat insecure position because no compelling reason has been found for many of the phenomena which occur. Nevertheless, a host of publications have appeared describing these effects, and attempting to give them some theoretical explanation. It soon became apparent that emulsification, the destruction of bacteria, the freeing of antibodies from bacterial cells, the acceleration of chemical

reactions, the production of luminescence and several other phenomena were dependent for their existence upon cavitation, which is the production of bubbles of gas in a liquid.[33] In gas-free liquids, or in liquids under high hydrostatic pressures, most of these disappear, and others, like emulsification, proceed only feebly and under restricted conditions. I have been told by every mathematical physicist I know that the analysis of cavitation is a task beyond the ability of present day mathematics. It is therefore necessary to examine the role of cavitation in the light of what experiment shows.

We may profitably fix our attention on the acceleration of chemical reactions because this usually requires a definite quantity of energy, which any proposed mechanism involving cavitation must be able to supply. The most common type of reaction to be reported is of the oxidation-reduction type, for which the production of iodine from sodium iodide may serve as an example. Since it is possible to measure the intensity of sound quantitatively,[34] it is possible to show by straightforward thermodynamical reasoning that no sound waves have ever been produced with sufficient pressure amplitude to affect



Jets of spray being sent out from a tube in which the walls are vibrating as stationary waves. Such vibration was produced by Wood and Loomis with high frequency sound waves of great intensity.



The propagation of ultrasonic energy through a diaphragm having a diameter equal to 8/10 wave-length. Picture on the left shows the waves while the one on the right shows the acoustic rays. (After E. Hiedemann and K. Osterhammel.)

a chemical reaction appreciably owing to the temperature difference between the compressed and rarefied portions of the waves. It is possible, perhaps, that the high temperatures developed at the surface of cavitated bubbles, which Marinesco has shown to reach 250°C,[35] may locally accelerate reactions to an extent which appears as a measurable acceleration of the reaction in the liquid as a whole. This might account for the acceleration of reactions which are not of the oxidation-reduction type, such as hydrolyses. But it has repeatedly been observed that the irradiation with intense sound of air-saturated water produces hydrogen peroxide if cavitation is permitted, but has no result if it is suppressed. Here we are not dealing with the hastening of a process which leads to equilibrium, but with the creation of a compound which requires the expenditure of some 3 electron volts, and no temperature rise due to bubble-collapse, bubble-heating and the like can bring it about by any mechanism known to me. It appears, however, that cavitation is accompanied by severe electrical disturbances. Just as the production of a spray of liquid by an "atomizer" causes a separation of charges, the production of a spray of holes, which is cavitation, has the same result. Mr. Hobart and I have shown, in a paper to be submitted to the *Journal of the American Chemical Society*, that large random electrical disturbances may be detected in a cavitating liquid. Over the distances comparable to molecular dimensions it appears reasonable to believe that these may be great enough to produce ionization, combination, and so on. Our proof is unfortunately not com-

plete, but it is clear that a ballo-electrical explanation of the role of cavitation in accelerating reactions is at present tenable. It also supplies a welcome explanation of the luminescence of liquids which Chambers[36] has studied.

There are many industrial applications of intense sound waves ranging from the destruction of fogs to the ageing of whiskey. Manufacturers of cosmetics and of photographic plates find intense sound useful in preparing stable and homogeneous emulsions. The milk people are interested both from the standpoint of homogenizing and of bacterial destruction. Paints and varnishes have had a crack at it. In fact, about 1932 there was a feeling in the air that anyone who manufactured anything, with the possible exception of horn buttons, was either installing a supersonic outfit or wishing that he had the money for one. The chief beneficiaries of this movement were the electric power companies. In some cases, it is true, better emulsions are produced by sound waves than by any other known method. The coagulation of fogs by airplanes zooming about with supersonic generators may have a great future. The smelting industry may find sound-waves indispensable for the degassing of metal melts. But the electrical production of sound waves is appallingly wasteful, and for the most part I believe that they will be supplanted by more efficient mechanical devices.

### *Bibliography*


1. Bourgin, Phil. Mag. **7**, 821 (1929).
2. Bourgin, Phys. Rev. **42**, 721 (1932).
3. Bourgin, Phys. Rev. **50**, 355 (1936).
4. Einstein, Sitz. Ber. Akad. 380 (1920).
5. Kneser, Ann. d. Physik **11**, 761 (1931); **16**, 337 (1933).
6. Rutgers, Ann. d. Physik **16**, 350 (1933).
7. Richards and Reid, J. Chem. Phys. **1**, 114 (1933).
8. Luck, Phys. Rev. **40**, 440 (1932).
9. Richards and Reid, J. Chem. Phys. **2**, 193 (1934).
10. Richards, J. Chem. Phys. **1**, 863 (1933).
11. Landau and Teller, Physik. Zeits. Sowjetunion **10**, 34 (1936).
12. Herzfeld and Rice, Phys. Rev. **31**, 691 (1928).
13. Richardson, Proc. Roy. Soc. **146A**, 56 (1934).
14. Penman, Proc. Phys. Soc. **47**, 543 (1935). The measurements described in this paper have been repeated by Warner, J. Acous. Soc. Am. **9**, 30 (1937), who has obtained results which are in accord with the collision theory within experimental error. Since Warner has preferred to present his measurements without interpretation I must take the responsibility for the statement that they refute Penman's arguments.
15. Fay, J. Acous. Soc. Am. **3**, 222 (1931).
16. Pielemeier, J. Acous. Soc. Am. **7**, 37 (1935).
17. Lucas, J. de phys. et rad. **8**, 41 (1937).
18. Pierce, Proc. Am. Acad. Sci. **63**, 1 (1928).
19. Grossmann, Physik. Zeits. **35**, 83 (1934).
20. Hubbard, Phys. Rev. **36**, 1688 (1931); **38**, 1011 (1931); **41**, 523 (1932); **46**, 525 (1934).
21. Debye and Sears, Proc. Nat. Acad. Sci. **18**, 409 (1932).
22. Brillouin, Ann. de physique **17**, 88 (1922).
23. Raman and Nath, Ind. Acad. Soc. Proc. **2A**, 406 (1935), and several subsequent papers.
24. Hubbard and Loomis, Phil. Mag. **5**, 1177 (1928).
25. Bär, Helv. Phys. Acta **8**, 500 (1935).
26. Hiedemann and Hoesch, Zeits. f. Physik **102**, 253 (1936). Nomoto, Phys. Math. Japan **19**, 255 (1937). Bergmann, Naturwiss. **25**, 113 (1937).
27. Knudsen, J. Acous. Soc. Am. **3**, 126 (1931); **5**, 112 (1933).
28. Knudsen and Obert, J. Acous. Soc. Am. **7**, 249 (1936).
29. Parthasarathy, Ind. Acad. Soc. Proc. **2A**, 497 (1935), and several subsequent papers.
30. Kneser and Knudsen, J. Acous. Soc. Am. **5**, 112, 122 (1933).
31. Eucken and Nümann, Zeits. f. physik. Chemie **36B**, 163 (1937), and earlier papers by Eucken there cited.
32. Wood and Loomis, Phil. Mag. **4**, 417 (1927).
33. Schmitt, Johnson, and Olson, J. Am. Chem. Soc. **51**, 370 (1929).
34. Richards, Proc. Nat. Acad. Sci. **17**, 611 (1931).
35. Marinesco, Comptes rendus **201**, 1187 (1935).
36. Chambers, J. Chem. Phys. **5**, 290 (1937).

This bibliography is only a skeleton. A complete list of publications on this subject would include several hundred titles.

# The Study of Extreme Pressures and their Importance in the Investigation of Engineering Problems

BY THOMAS C. POULTER

*Executive Director, Research Foundation of Armour Institute of Technology, Chicago, Illinois*

PERHAPS the earliest high pressure work of note is the celebrated attempt of the Florentine Academy to find whether water is compressible. Using a lead container in which they probably had much less than a thousand pounds pressure, they concluded that water was an incompressible liquid. In 1762–64 Candon carried out some experiments in which he proved that water was compressible, and in spite of the fact that water is sometimes referred to today as an incompressible liquid, experiments have been carried out in which water has been compressed to one-half its normal value. The fifty-year period ending 1869 was one in which a great deal was accomplished in the field of high pressure in spite of the fact that very few good numerical results were obtained. Since then a great deal of activity has been shown in the field of extreme pressure investigation, and the literature contains hundreds of articles published by a great many investigators so that at the present time there is a wealth of high pressure data available.

If we examine the curve which represents the advance that has been made in the maximum pressures employed in this work in the past two hundred years, Fig. 1, we find that it is rising at such an increasing rate that it makes even the high pressure investigator wonder what the future may lead to in this field. In spite of the fact that man has until recently been unable to develop pressures as high as one million pounds per square inch, such pressures are frequently occurring about us in our everyday experiences. They are, of course, confined to an extremely small space, and have a very short duration. Pressures of a few hundred thousand pounds per square inch are occasionally produced in the hypoid gear in the normal operation of an automobile, in roller or ball bearings, under the wheel of a glass cutter, the bullet of a high powered rifle striking a solid object, and in many other cases where two hard objects strike each other. Such extreme liquid or gas pressures are of course much less common, but even gas pressures of nearly a million pounds per square inch are probably occurring every day as particles of meteoric matter traveling with a velocity of thirty miles per second, collide with the very light upper layer of the earth's atmosphere.

Recently an experiment was carried out in the high pressure laboratory of the Research Foundation of Armour Institute of Technology in which a pressure of one million five hundred thousand pounds per square inch was developed. This was the highest pressure ever produced experimentally in high pressure investigations. The pressure at the center of the earth is estimated to be about three million two hundred thousand atmospheres, or only thirty-two times as great as the maximum pressure that has been produced in the laboratory. It is, of course, experimentally impossible to produce in the laboratory pressures corresponding to those at the center of the earth, but this increase in range of pressures may clear up many puzzling questions which have formerly been considered impossible of solution. Very little attempt has been

made in the past to utilize extreme pressure experiments in the solution of everyday problems, but they are rapidly assuming a place of increasing importance, and in such experiments lie the answers to many questions of greatest commercial importance.

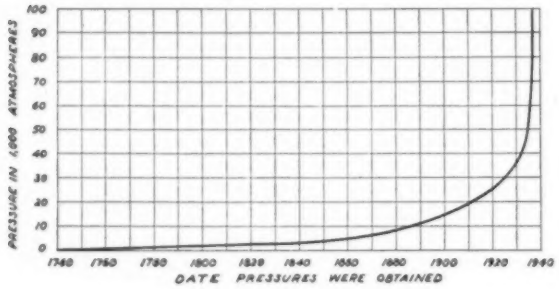



FIG. 1.

A brief survey of the work that has been studied by means of the high pressure equipment now in the high pressure laboratory of the Research Foundation of Armour Institute of Technology will serve to show the great diversified application of such investigations. The work was started in an attempt to study a reported effect of pressure upon the chemical reaction of sulphuric acid on zinc to form zinc sulphate and liberate hydrogen.[1] This investigation was carried to a pressure of thirty thousand atmospheres, during which it was found that the effect we set out to study did not exist. In the course of the investigation we found many interesting things including a little known chemical reaction of hydrogen on sulphuric acid to give hydrogen sulfide and water. Other chemical reactions were investigated under extreme pressures, such as the hydrolysis of sucrose. The rate of this reaction was found to decrease with increase in pressure. From the information gained as a result of our previous work, new pressure equipment was developed in which it was possible to carry on experiments in cylinders with glass, quartz or diamond windows,[2–4] which will withstand pressures of one-half million pounds per square inch. In connection with this development many interesting phenomena were observed, such as the ability of glass or quartz windows one-quarter inch thick to be bent to a radius of curvature of four inches as many as ten times before being broken. The development of such windows made it possible to study the effect of pressures upon the optical rotation of optically active compounds, which effect was to increase the rotation of all compounds studied, some of them being increased to threefold their normal value. Both dextro- and levorotatory compounds were investigated. The windows also made it possible to study the phosphorescence of zinc sulfide,[5] which effect is to decrease the intensity of the phosphorescence by a factor of one-half for a pressure of thirty thousand atmospheres. It was also found that a rapid change in pressure would produce a bright glow of the zinc sulfide. Aside from the intensity, the fluorescent and phosphorescent properties of zinc sulfide were very little affected by extremely high pressures. It, therefore, provided a convenient method of studying the effect of extreme pressures upon the radioactive decomposition[5] of various materials by actually observing scintillation. These measurements confirmed the negative results of similar investigations by other observers using quite different methods. Further investigations made possible by the pressure windows were the determination of the effect of pressure upon the index of refraction[6] of certain liquids. This investigation showed that the Lorenz-Lorentz relation holds for those substances investigated. The quartz windows were further utilized in making absorption spectra measurements of the effect of pressure upon the compressibility of the neodymium atoms at the various electron energy levels. The effect of pressure upon living organisms has been studied, with the observation that twelve thousand atmospheres are necessary to kill bacteria, but as the complexity of the living organism increases the pressure necessary to destroy life decreases. Hydra and planaria were found to withstand pressures of from ten to twenty thousand pounds per square inch without any serious damage. The effect of high pressures was to precipitate some of the colloidal constituents of the organism, and a study was made of the precipitation of other colloids,[7] such as sulfur, silver, gold, ferric hydroxide, molybdenum blue, and prussian blue. Colloidal ferric hydroxide is precipitated completely by pressures as low as 100 atmospheres whereas molybdenum blue is only slightly precipitated at a pressure of 17,000 atmospheres.

The electrode potential[1] of the hydrogen electrode was investigated under pressures up to thirty thousand atmospheres and the e.m.f. of the Weston Standard Cell[8] up to twelve thousand atmospheres. The effect of pressure on both of these is of a rather low order of magnetism. The effect of pressure upon the penetration of water and numerous other liquids into glass and metal surfaces[9] has been investigated with some very interesting results. Water was found to penetrate in considerable quantities to a depth of several millimeters into glass in only a few minutes time. The penetration of alcohol and ether is somewhat less, and such liquids as paraffin oils, glycerin, etc. penetrate scarcely at all. A similar effect, but to a considerably lesser degree, was observed for the penetration of liquids into metals. The penetration of gases into metals[9] of course is more rapid than that of liquid, and represents a problem of considerable industrial importance, particularly in the case of penetration of hydrogen into steel[10] whereby the tensile strength of the steel is reduced to less than half its original value. The normal dissociation of steam into hydrogen and oxygen, and the subsequent removal of the oxygen by its reaction with the metal of the high pressure steam lines, particularly under conditions where a high superheat is used, presents a problem of utmost importance in the high pressure high superheat steam installations that are being used at the present time. It has been shown that considerable quantities of hydrogen are continually escaping through the walls of the high pressure high superheat steam lines, and this subject should be thoroughly investigated to determine to what extent this is affecting the tensile strength of the metal of the steam lines. If this effect is cumulative, as is the case under the conditions of many of our experiments, it might very well become a problem of utmost importance in high superheat steam insulation, if not indeed a limiting factor in the industry. We have found that this effect is not only a function of the temperature and pressure, but is also affected by the composition and heat treatment of the steel. It is, therefore, a combination high pressure and metallurgical investigation.

Other fields which have been investigated are: compressibility of liquids[4] and solids, the effect of pressure upon the viscosity of liquids, and in some cases the change of state, or even polymorphic transition.[1–6]

A brief summary of the effect of pressure upon the properties of matter will reveal some very interesting facts. The following characteristics for liquids and solids are all affected by pressure: density, volume, index of refraction, electrical conductivity, thermal conductivity, magnetic permeability, dielectric constant, optical rotation, chemical reactivity, solubility, phosphorescence, fluorescence, physical strength, specific heat, latent heat, permeability to gases and liquids and viscosity. If we consider all but the last two items on this list, we will see that the effect is anywhere from a small percent up to certainly less than a factor of ten. But now let us consider the last two items. The permeability to gases and liquids may be of the order of magnitude of many thousand-fold, and the coefficient of viscosity of many lubricating oils is increased by as much as two hundred million times their value at atmospheric pressure. It is, therefore, our belief that the permeability to gases and liquids and viscosity are the two characteristics that will represent the most fruitful problems for investigation and certainly are the two that are most likely to produce noticeable effects in industrial processes involving high pressures. Investigation of the permeability of gases and

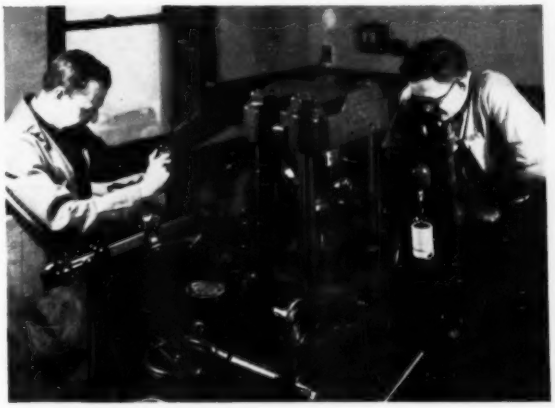


FIG. 2. Set-up used for absorption spectra measurements of materials under pressures of as high as 580,000 pounds per square inch.

liquids has already been discussed and the viscosity investigation of lubricating oils under pressure probably represents the most fruitful investigations to be carried out in this field.

FIG. 3. Assembly used for producing a pressure of 1,500,000 pounds per square inch.

Anyone familiar with tests on lubricating oils is aware of the fact that a small quantity of suspended solids in the oil is very likely to produce a bearing failure even though these particles may be softer than the two metals being lubricated. This effect is in part, if not largely, the result of the solid materials which may not be so soft at the higher pressures, so completely wiping off the lubricant as to permit the two metal surfaces to come in contact and start the failure. It is the opinion of the author that many failures under conditions where an extreme pressure lubricant is required are caused by a similar phenomena, but that it has not been generally recognized because an examination of the oil before failure shows no such foreign particles, and after the failure only such particles as have been removed from the metal surface during failure. One very likely source of such suspended particles has in general been overlooked, namely those which are liquids at normal pressures but which separate as solids under the extreme pressure. In some cases, these solids may be of metallic hardness. Water in the form of an emulsion, or even the normal small amount of moisture in the oil is converted to a solid at one-third to one-fourth the pressure frequently existing in an oil where an extreme pressure lubricant is required. The solid thus formed is known as ice VI and is very much harder than ordinary ice. Under these pressures it has a melting point about as high as the normal boiling point of water.

Some of the normal constituents of certain oils may go through a similar transition, producing a substance hard enough to wipe the lubricant from the metal surfaces, or even produce some abrasion action itself. Another phenomenon which has frequently been observed in our work at ultra pressures and which no doubt has its effect not only upon the wear, but upon subsequent failure, is the actual penetration of water or other compressed materials into the surface of the steel. The escape of such substances when the pressure is released causes a breakdown of the metal surface.[9] If a glass rod in contact with water is subjected to such pressures and the pressure released immediately, the rod will not be affected, but if the pressure is maintained for five minutes, so as to allow time for the water to penetrate the surface of the glass, and then the pressure is released, the rod will be broken up into disks; whereas, if the pressure is maintained for twenty minutes permitting the water to penetrate deeply into the glass, then released rapidly, the glass will be shattered into small pieces. Similar effects, although to a much lesser degree, are known for steel and other metals.

With the increasing realization that pressures of as high as four hundred thousand pounds per square inch exist between two metal surfaces which are supposed to be kept separated by means of a lubricant, has led to an entirely different conception of the problem. Extreme pressure lubrication is only in its infancy, and a proper method of approach to a fundamental understanding of the problem is through research at extreme pressures.

Suppose we put a lubricating oil under such a pressure, and study its physical properties. We find that it is no longer a fluid in the normal

sense of the term, but that it has in some cases become quite hard. It is indeed fortunate that this is true, for if such were not the case, it would be impossible to maintain a film of lubricant between many of the operating parts of a modern automobile.

The viscosity of a liquid at normal atmospheric pressure is no indication of what its viscosity or hardness may be at greatly elevated pressures. Two liquids having almost identical physical properties at one pressure may vary by several thousand-fold at another pressure well within the range of those normally encountered in extreme pressure lubrication problems. The effect of pressure upon the physical properties of various liquids differs so widely as to make it possible to produce a mixture of almost any desired viscosity or hardness at almost any specified pressure.

It is the opinion of the author that the day is not far distant when the specifications of a lubricant will be largely based upon the pressure hardness curves for not only the mixture as a whole, but upon the individual constituents. This is particularly true in view of the modern conception that a lubricant is a fluid without harmful properties serving as a carrier for a very small quantity of material which produces a lubricating action. The proper hardness curves can be theoretically determined from the condition under which the oil is intended to operate, taking into account the hardness and curvature of the two surfaces, the rate of motion of the contact point with respect to these two surfaces, the rate of motion of one surface with respect to the other, and other conditions affecting the plastic flow of the lubricant. With these specifications available, it will be possible for the lubrication engineer to design a lubricant that will have the physical characteristics necessary for it to serve as a lubricant in the true sense of the word, and which will prevent abrasion of the two metal surfaces.

It was the development of a new steel that permitted the production of an automobile requiring an E P lubricant. One has but to consider the present trends in the development of alloy steels to realize that the time is already at hand when an E P lubricant of a still higher order of magnitude is required.

### *Bibliography*


1. Thomas C. Poulter and Glen E. Frazer, "A Study of the Action of Acids on Zinc at Pressures of from One to Thirty Thousand Atmospheres," J. Phys. Chem. **38**, 1131–1140 (1934).
2. Thomas C. Poulter, "Apparatus for Optical Studies at High Pressures," Phys. Rev. **40**, 860–871 (1932).
3. Thomas C. Poulter and Carl Benz, "Lenz Effect of Pressure Windows," Phys. Rev. **40**, 872–876 (1932).
4. Thomas C. Poulter and Francis Buckley, "Diamond Windows for Withstanding Very High Pressures," Phys. Rev. **41**, 364–365 (1932).
5. Thomas C. Poulter and Harold McComb, "A Study of Phosphorescent Zinc Sulphide Screens and Radioactivity Under Extremely High Pressure," Iowa Academy of Science, Abstracts, **37**, 311–312 (1930).
6. Thomas C. Poulter, Carter Ritchey and Carl A. Benz, "Effect of Pressure on the Index of Refraction of Paraffin Oil and Glycerine," Phys. Rev. **41**, 366–367 (1932).
7. Robert Wilson and Thomas C. Poulter, Iowa Academy of Science, **36**, 295–296 (1929).
8. Thomas C. Poulter and Carter Ritchey, "Effect of Pressure on the e.m.f. of the Weston Standard Cell," Phys. Rev. **39**, 816–820 (1932).
9. Thomas C. Poulter and Robert O. Wilson, "Permeability of Glass and Fused Quartz to Ether, Alcohol, and Water at High Pressure," Phys. Rev. **40**, 877–880 (1932).
10. Thomas C. Poulter and Lester Uffelman, "The Penetration of Hydrogen Through Steel at Four Thousand Atmospheres," Physics **3**, 147–148 (1932).

# Application of the Lagrangian Equations to Electrical Circuits

BY D. A. WELLS

*University of Cincinnati, Cincinnati, Ohio*

IT has been known for many years that the Lagrangian equations of analytical dynamics can be applied to electrical circuits as a means of arriving at the differential equations of current flow and mechanical motion of moving parts of the circuit. However, this method has never come into general use even though it is quite easy to apply and, for certain types of problems to be illustrated later, is a remarkably powerful tool. There seem to be two reasons why the Lagrangian method has not come to be used more generally. First, most simple circuits with no moving parts can be handled more quickly by a direct application of Kirchhoff's laws and it is a fact that most circuits are of this type. Secondly, the textbooks of applied electricity almost invariably omit a discussion of this method. Hence the electrical engineer has, in most cases, not had an opportunity to become familiar with the use of this powerful tool.

The purpose of this paper is, therefore, to discuss briefly a convenient form of the Lagrangian equations applicable to electrical circuits and to illustrate their use with a number of specific examples. A brief review of the application of Lagrange's equations to mechanical systems will first be given. The derivation of these equations will not be repeated here since it may be found in any advanced text on analytical dynamics.

If the generalized coordinates necessary to represent the configuration of a holonomic mechanical system are denoted by, $q_1, q_2, \cdots, q_n$, the Lagrangian equations of motion of the system may be written as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_r}\right)-\frac{\partial L}{\partial q_r}=F_r \qquad (1)$$

where the symbols have the following meaning: $q_r$ represents any one of the generalized coordinates, $\dot{q}_r$ is the first time derivative of $q_r$, and $L$, the so-called "Lagrangian function," is the kinetic energy $T$ minus the potential energy $V$ of the system. That is $L=T-V$. $F_r$ represents a generalized dissipative force such that $F_r\delta q_r$ gives the energy dissipated by the system when $q_r$ is changed to $q_r+\delta q_r$, all other coordinates remaining unchanged. Since (1) is true for any one of the coordinates, $q_1, q_2, \cdots, q_n$, it represents a set of $n$ second order differential equations which when solved give the complete motion of the system.

As an illustrative example, consider the motion of the double pendulum shown in Fig. 1. Let it be assumed that the motion is confined to a plane, that the lengths $r_1$ and $r_2$ are fixed and that a damping vane is attached to $m_1$. It is seen that under these conditions only two coordinates are required to determine completely the positions of the masses $m_1$ and $m_2$. Of the many choices of coordinates that could be made, the angles $\theta$ and $\varphi$ are perhaps the most convenient. We shall now set up the Lagrangian function $L$ and by an application of (1) write out in explicit form the two differential equations of motion of the pendulum.

The kinetic energy of $m_1$ is seen to be

$$T_1=(m_1/2)(r_1\dot{\theta})^2.$$

The kinetic energy of $m_2$ can most easily be found by first writing down its position in the $x$ and $y$ coordinates. From the diagram it

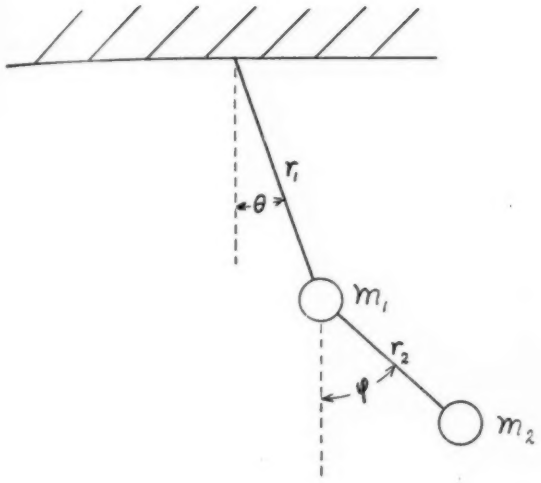



FIG. 1. A double pendulum for illustrating the application of Lagrange's equations to particle dynamics.

follows that

$$x = r_1 \sin\theta + r_2 \sin\varphi + x_0,$$
$$y = r_1(1-\cos\theta) + r_2(1-\cos\varphi) + y_0,$$

where $x_0$ and $y_0$ are the coordinates of $m_2$ in its rest position. Hence, the $x$ and $y$ components of velocity are

$$\dot{x} = r_1\dot{\theta}\cos\theta + r_2\dot{\varphi}\cos\varphi,$$
$$\dot{y} = r_1\dot{\theta}\sin\theta + r_2\dot{\varphi}\sin\varphi.$$

The kinetic energy of $m_2$ is, therefore,

$$T_2 = (m_2/2)(\dot{x}^2+\dot{y}^2)$$
$$= (m_2/2)[r_1^2\dot{\theta}^2 + r_2^2\dot{\varphi}^2 + 2r_1r_2\dot{\theta}\dot{\varphi}\cos(\varphi-\theta)]$$

and the total kinetic energy is given by

$$T = T_1 + T_2 = \tfrac{1}{2}[(m_1+m_2)r_1^2\dot{\theta}^2 + m_2r_2^2\dot{\varphi}^2 + 2m_2r_1r_2\dot{\theta}\dot{\varphi}\cos(\varphi-\theta)].$$

The potential energy of the system, referred to the lowest possible position of $m_2$, is easily seen to be

$$V = (m_1+m_2)gr_1(1-\cos\theta) + m_2gr_2(1-\cos\varphi).$$

Assuming that the viscous drag on the vane attached to $m_1$ is proportional to its linear velocity (the proportionality factor being $k$), the change in energy of the system by a small displacement, $r_1\delta\theta$, of $m_1$ is given by

$$\delta\mathcal{E} = -(kr_1\dot{\theta})(r_1\delta\theta) = F_1\delta\theta.$$

Hence the generalized dissipative force $F_1$ corresponding to the coordinate $\theta$ is

$$F_1 = -kr_1^2\dot{\theta}.$$

Since the damping of $m_2$ is assumed to be negligible, $F_2 = 0$. The Lagrangian function is now seen to be

$$L = \tfrac{1}{2}[(m_1+m_2)r_1^2\dot{\theta}^2 + m_2r_2^2\dot{\varphi}^2 + 2m_2r_1r_2\dot{\theta}\dot{\varphi}\cos(\varphi-\theta)] - (m_1+m_2)gr_1(1-\cos\theta) - m_2gr_2(1-\cos\varphi).$$

From which,

$$(\partial L/\partial\dot{\theta}) = (m_1+m_2)r_1^2\dot{\theta} + m_2r_1r_2\dot{\varphi}\cos(\varphi-\theta),$$
$$(d/dt)(\partial L/\partial\dot{\theta}) = (m_1+m_2)r_1^2\ddot{\theta} + m_2r_1r_2\ddot{\varphi}\cos(\varphi-\theta) + m_2r_1r_2\dot{\varphi}(\dot{\theta}-\dot{\varphi})\sin(\varphi-\theta),$$
$$(\partial L/\partial\theta) = m_2r_1r_2\dot{\theta}\dot{\varphi}\sin(\varphi-\theta) - (m_1+m_2)gr_1\sin\theta.$$

Therefore, by (1), the first equation of motion is

$$(m_1+m_2)r_1^2\ddot{\theta} + m_2r_1r_2\ddot{\varphi}\cos(\varphi-\theta) + m_2r_1r_2\dot{\varphi}(\dot{\theta}-\dot{\varphi})\sin(\varphi-\theta) - m_2r_1r_2\dot{\theta}\dot{\varphi}\sin(\varphi-\theta) + (m_1+m_2)gr_1\sin\theta = -kr_1^2\dot{\theta}.$$

Likewise the second equation of motion is found to be

$$m_2r_2^2\ddot{\varphi} + m_2r_1r_2\ddot{\theta}\cos(\varphi-\theta) + m_2r_1r_2\dot{\theta}(\dot{\theta}-\dot{\varphi})\sin(\varphi-\theta) + m_2r_1r_2\dot{\theta}\dot{\varphi}\sin(\varphi-\theta) + m_2gr_2\sin\varphi = 0.$$

When $L$ and the dissipative forces $F_r$ have been properly determined, Lagrange's equations give the differential equations of motion of *any mechanical system in any system of coordinates.* (It should be noted that when $L$ and $F_r$ have been determined the labor of determining the equations of motion is over. The final step or the application of (1) amounts only to a bit of simple differentiation.)

Likewise it is possible to find a Lagrangian function and dissipative "forces" such that

Lagrange's equations will give the differential equations of an electrical circuit, *however complex, even if various parts of the circuit are allowed to move in any way with respect to each other.*

### *The Lagrangian Function for an Electrical Circuit*

If (1) is to be used as a means of arriving at the differential equations of an electrical circuit, it follows that we must find a suitable Lagrangian function $L$ and the proper dissipative forces $F_r$. This implies a choice of "coordinates" and a determination of the "kinetic" and "potential" energy of the system. To begin with, it will be assumed that no part of the circuit is allowed to move. Thus "mechanical" kinetic and potential energies will not enter the first part of the discussion.

An explanation of the choice of coordinates here made cannot be given in a few words, but it will become clear from what follows that the charges which, from some initial time $t_0$, have passed through the various branches of the circuit constitute an appropriate set of coordinates. Hence the charges, represented by $Q_1$, $Q_2$, $\cdots$, $Q_n$, correspond in an electrical circuit to the coordinates $q_1$, $q_2$, $\cdots$, $q_n$, in a mechanical system. A fact of great importance in the use of charges as generalized coordinates is that there are almost invariably relations between certain of the charges. That is, not all of the coordinates are independent. At any junction in a circuit the algebraic sum of the charges flowing to the junction must be zero (Kirchhoff's law). Hence an application of this law gives a set of equations, by means of which, certain coordinates (charges) can be eliminated. For example, an unbalanced Wheatstone bridge circuit contains six branches and therefore there are six currents and six corresponding charges. However, there are three independent relations between these charges and hence, by elimination, the six can be reduced to three. It is a matter of indifference as to which three are eliminated. In the language of analytical dynamics, the Wheatstone bridge is a problem of three "degrees of freedom"—not six as might first be supposed. By this method an electrical circuit *must be reduced to its proper number of degrees of freedom.* If this is not done, Lagrange's equations, in the form in which we are to use them, will not apply. At the outset of work on a circuit each current and each electromotive force must be assigned a direction just as if the circuit were to be handled by the usual methods. This makes it possible to determine the proper number of degrees of freedom and, as will be seen later, the potential energy of the sources.

A consideration of the magnetic energy of an inductance will show, at once, what corresponds to kinetic energy in the electrical circuit. The energy of the magnetic field about a coil of self-inductance $M$, is $MI^2/2$ where $I=(dQ/dt)=\dot{Q}$ is the current in the coil. Since $Q$ is considered as a coordinate, $\dot{Q}$ is a generalized velocity and it is seen that $M\dot{Q}^2/2$ has the same form as the familiar term $mv^2/2$ representing the kinetic energy of a particle.

Consider two coils having self-inductances of $M_1$ and $M_2$ and each a mutual inductance with respect to the other of $M_{12}$. Denoting the current through the first by $\dot{Q}_1$ and through the second by $\dot{Q}_2$, the energy of inductance is given by

$$\mathcal{E}=\frac{M_1\dot{Q}_1^2}{2}+\frac{M_2\dot{Q}_2^2}{2}+M_{12}\dot{Q}_1\dot{Q}_2.$$

But this is again a quadratic function of the generalized velocities and corresponds closely in form to the expression for the kinetic energy of the double pendulum. A further consideration of the energy of inductance shows that it can always be expressed as a quadratic function of the currents or generalized velocities. Hence, *energy of inductance is to be regarded as kinetic energy.*

The potential energy of a circuit is made up of two parts, the energy of the sources (batteries etc.), and the energy stored in condensers that may be connected in the circuit. A source will always deliver, or tend to deliver, charge to the circuit. If a charge $Q$ is delivered by a source having a constant electromotive force $E$, the source has less energy to the extent of $EQ$. Regarding the source as having potential energy and referring it to the "point" $Q=0$, it is seen that this potential energy must be written as $-EQ$. This amounts to saying that in order to reduce a certain delivered charge to zero, work must be done *on the source.* Of course, if a charge

is regarded as flowing against the electromotive force of the source, the potential energy $EQ$ must be considered positive. That the potential energy of a source with constant electromotive force is exactly analogous to the potential energy of a particle of mass $m$ acted upon by gravity, is easily seen. If $y$, Fig. 2, is considered positive down, the potential energy of $m$ is $-mgy$, $mg$ corresponding to $E$ and $y$ to $Q$. It is important to note that even though the electromotive force of a source may vary with time, as for example $E=E_0 \sin \omega t$, $-EQ$ is still to be regarded as a potential energy. This is true for the reason that, in the derivation of Lagrange's equations, time is not allowed to vary as any one of the coordinates $q_r$ is changed to $q_r+\delta q_r$.[1]

The energy of a charged condenser is $Q^2/2C$, where $C$ is the capacity. Since this expression has exactly the same form as the expression for the potential energy of a stretched spring, it is immediately suggested that the energy of a charged condenser be regarded as potential energy. It is clear that this energy must be taken as positive since, when the charge is reduced to zero, energy is expended by the condenser—not on it. Hence the potential energy of a circuit is the algebraic sum of the energies of the sources plus the sum of the energies of the condensers.

It only remains to determine what is to be regarded as the generalized dissipative forces. Neglecting radiation, hysteresis, and eddy currents, the electrical energy of a circuit can only be dissipated as heat in the resistances. Of course, electrical energy can be changed to mechanical energy by allowing parts of the circuit to move but, as previously mentioned, we are for the moment assuming that each part of the electrical system is fixed. The rate of change of energy of a circuit containing only one resistance $R$, due to the passage of current through this resistance, is given by

$$(d\epsilon/dt) = -R\dot{Q}^2.$$

But $F\dot{Q}$, where $F$ is the generalized force and $\dot{Q}$ the current through the resistance or generalized velocity, must also represent the rate of change of energy of the system. Hence,

$$F\dot{Q} = -R\dot{Q}^2.$$

Thus the generalized dissipative force is determined by

$$F = -R\dot{Q}.$$

A complex circuit requires as many differential equations for its treatment as there are degrees

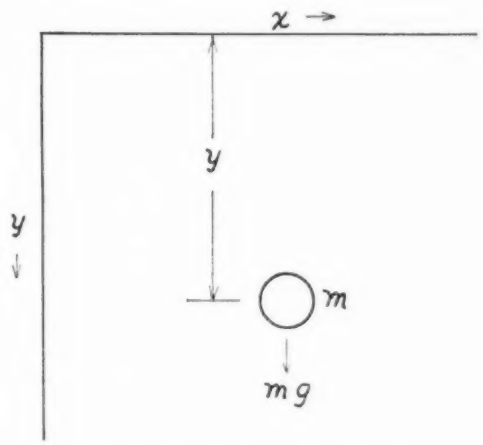



FIG. 2. Simple diagram to illustrate the analogy between gravitational potential energy and the potential energy of an electrical source.

of freedom (independent currents or charges). Each equation involves a generalized dissipative force and as a general rule each force involves a number of currents and resistances. The method of arriving at dissipative forces for a complex circuit is quite easy. It is best shown by the examples that follow.

We, therefore, write the Lagrangian equations applicable to a stationary electrical circuit as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}_r}\right) - \frac{\partial L}{\partial Q_r} = F_r, \qquad (2)$$

where $Q_r$ represents any one of the $n$ charges necessary for the determination of all other charges of the circuit, $\dot{Q}_r$ the corresponding current, $F_r$ the corresponding electrical dissipative force and where, in the Lagrangian function $L=T-V$, the kinetic energy $T$ and potential energy $V$ are determined as outlined above.

In order to make clear just how the $L$ function is set up and how the circuit equations are ob-

[1] See Whittaker, *Analytical Dynamics*, p. 36.

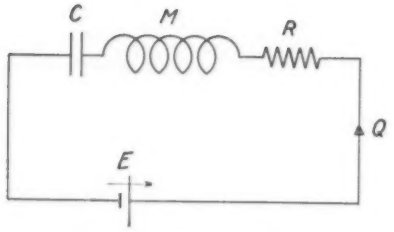



FIG. 3. Simple series circuit.

tained from the Lagrangian equations, two specific examples will now be given. Consider first the simple series circuit of Fig. 3. The kinetic energy $T=M\dot{Q}^2/2$ and the potential energy $V=-EQ+Q^2/2C$. Hence the Lagrangian function appropriate to this circuit is

$$L=T-V=\frac{M\dot{Q}^2}{2}+EQ-\frac{Q^2}{2C}.$$

The dissipative force $F=-R\dot{Q}$. Therefore, since

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}}\right)=M\ddot{Q} \quad \text{and} \quad \frac{\partial L}{\partial Q}=E-\frac{Q}{C},$$

an application of (2) gives

$$M\ddot{Q}-E+Q/C=-R\dot{Q},$$

which is the familiar differential equation of this circuit.

Consider next the simple circuit of Fig. 4. We shall assume that the inductance coils each have a mutual inductance of $M_{12}$ for the other. The charges to be considered are $Q$, $Q_1$, and $Q_2$. But since $Q=Q_1+Q_2$, one of these charges, say $Q$, can be eliminated. Hence the kinetic and potential energies are given by

$$T=\frac{M_1\dot{Q}_1^2}{2}+\frac{M_2\dot{Q}_2^2}{2}+M_{12}\dot{Q}_1\dot{Q}_2,$$

$$V=-E_1(Q_1+Q_2)+E_2Q_2+\frac{Q_1^2}{2C_1}+\frac{Q_2^2}{2C_2}.$$

Since this is a problem of two degrees of freedom, the coordinates being $Q_1$ and $Q_2$, there are two corresponding generalized dissipative forces $F_1$ and $F_2$ which can easily be found as follows. Suppose the charge $Q$ is thought of as increasing slightly. The change in energy of the system is given by

$$\begin{aligned}\delta\mathcal{E}&=-R_1\dot{Q}_1\delta Q_1-R_2\dot{Q}_2\delta Q_2-R_3(\dot{Q}_1+\dot{Q}_2)(\delta Q_1+\delta Q_2)\\&=-[R_1\dot{Q}_1+R_3(\dot{Q}_1+\dot{Q}_2)]\delta Q_1\\&\qquad-[R_2\dot{Q}_2+R_3(\dot{Q}_1+\dot{Q}_2)]\delta Q_2.\end{aligned}$$

Hence, for a slight change of $Q_1$ only,

$$\delta\mathcal{E}_1=-[R_1\dot{Q}_1+R_3(\dot{Q}_1+\dot{Q}_2)]\delta Q_1.$$

Setting this equal to the generalized force $F_1$ times the displacement $\delta Q_1$ we get

$$F_1=-[R_1\dot{Q}_1+R_3(\dot{Q}_1+\dot{Q}_2)].$$

Likewise $F_2=-[R_2\dot{Q}_2+R_3(\dot{Q}_1+\dot{Q}_2)]$.

Thus, from $$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}_1}\right)-\frac{\partial L}{\partial Q_1}=F_1$$

and $$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}_2}\right)-\frac{\partial L}{\partial Q_2}=F_2,$$

we obtain the two differential equations of "motion" of the circuit

$$\begin{aligned}M_1\ddot{Q}_1+M_{12}\ddot{Q}_2-E_1+(Q_1/C_1)&\\=-R_1\dot{Q}_1-R_3(\dot{Q}_1+\dot{Q}_2),&\\M_2\ddot{Q}_2+M_{12}\ddot{Q}_1-E_1+E_2+(Q_2/C_2)&\\=-R_2\dot{Q}_2-R_3(\dot{Q}_1+\dot{Q}_2).&\end{aligned}$$

These differential equations and the one for the first example are, of course, just the equations usually obtained by an application of Kirchhoff's laws. It is seen that the Lagrangian method has little if any advantage when applied to the simple type of circuit illustrated thus far. The real power and usefulness of the method becomes evident when applied to circuits where certain parts are allowed to move. Following a brief discussion of the form taken by the Lagrangian function for such a circuit, two further examples will be given.

If parts of an electrical circuit are in motion, the total kinetic energy of the system must include the kinetic energy of motion of the parts as well as the energy of inductance. Likewise the expression for the potential energy must include terms for the potential energy that the circuit, treated as a mechanical system, may have as

well as the potential energy of the sources and energy of the condensers. In other words, *the Lagrangian function must contain the kinetic and potential energies of a mechanical system as well as the electrical circuit.* This means that the Lagrangian contains two groups of generalized coordinates, the charges $Q_1, Q_2, \cdots, Q_n$, and the usual displacement coordinates $q_1, q_2, \cdots, q_m$. The number of differential equations now obtained and the number of generalized dissipative forces required will be $n+m$. If there are no forces of friction, viscous drag, etc., acting on the moving parts the generalized forces corresponding to $q_1, q_2, \cdots, q_m$, are each zero.

As a simple example of an electrical system containing moving parts consider the circuit shown in Fig. 5. The upper plate of the condenser $C$, having a mass of $m$, is attached to the coil spring and is free to move in a vertical direction under the action of the spring and electrical forces. Gravity is not considered. All other parts of the circuit are fixed. The unusual features of this circuit are the variable capacity $C$, the kinetic energy of the moving plate, and the potential energy of the spring.

Call $s$ the distance between condenser plates when the spring is unstretched and $x$ the displacement of the plate from the position it occupies when the spring is unstretched. The potential energy of the spring is thus $kx^2/2$, where $k$ is the usual spring constant. The capacity of the condenser is given by

$$C=\frac{A}{4\pi(s+x)}=\frac{a}{s+x},$$

where $A$ is the area of one plate and the dielectric constant is taken as unity. The potential energy of the condenser is, therefore, $(Q^2/2a)(s+x)$ and the kinetic energy of the moving plate is $m\dot{x}^2/2$. Thus the Lagrangian for this system is seen to be

$$L=\frac{M\dot{Q}^2}{2}+\frac{m\dot{x}^2}{2}+EQ-\frac{Q^2}{2a}(s+x)-\frac{kx^2}{2}.$$

This is a problem of two degrees of freedom, the two generalized coordinates being $Q$ and $x$. From the general Lagrangian equation the two differential equations of the circuit follow at once:

$$M\ddot{Q}+(Q/a)(s+x)-E=R\dot{Q},$$

$$m\ddot{x}-Q^2/2a+kx=0.$$

A solution of these two simultaneous differential equations gives the mechanical behavior of the moving plate as well as the electrical behavior of the circuit. If a mechanical dissipative force is allowed to act on the moving plate it can easily be taken into account. For example, if a viscous drag (given by $f=-b\dot{x}$) is acting on the moving plate, the second equation becomes

$$m\ddot{x}-Q^2/2a+kx=-b\dot{x}.$$

In the more general problem any part of the circuit may be allowed to move. As a final example and one which enables us to write down the Lagrangian function in a very general way, consider the circuit of Fig. 6. Due to the length of details that would be involved several steps in the work are left in implicit form. It is believed, however, that no details are omitted that would cause great difficulty in handling an actual problem.

There are four charges to be considered but they are not all independent. It is seen from the

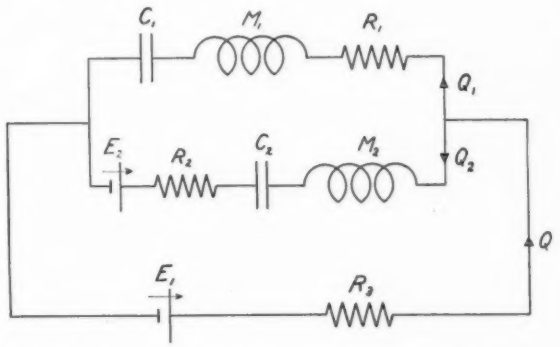



FIG. 4. Circuit to illustrate the method of obtaining "dissipative forces."

diagram, assuming that the charges are flowing in the directions indicated, that $Q_3-Q_4=Q_2-Q_1$. Hence any one of the charges can be eliminated and the problem, so far as the electrical part is concerned, is one of three degrees of freedom. In the relations that follow we will eliminate $Q_4$.

Assuming that the coils each have the self-inductance indicated on the diagram and that each has a mutual inductance with respect to

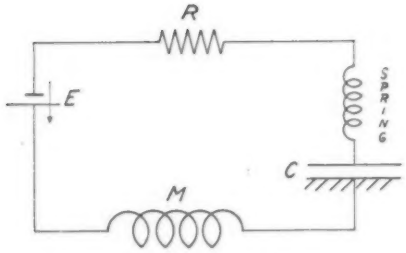



FIG. 5. Circuit containing a moving plate condenser.

every other, the energy of inductance or the energy corresponding to kinetic energy is given by

$$T_{\text{elect.}}=\tfrac{1}{2}(M_{11}\dot{Q}_1^2+M_{12}\dot{Q}_1\dot{Q}_2+M_{13}\dot{Q}_1\dot{Q}_3)+\tfrac{1}{2}(M_{21}\dot{Q}_2\dot{Q}_1+M_{22}\dot{Q}_2^2+M_{23}\dot{Q}_2\dot{Q}_3)+\tfrac{1}{2}(M_{31}\dot{Q}_3\dot{Q}_1+M_{32}\dot{Q}_3\dot{Q}_2+M_{33}\dot{Q}_3^2),$$

where, for example, $M_{23}$ is the mutual inductance of the second and third coils. This can be written in compact form as

$$T_{\text{elect.}}=\tfrac{1}{2}\sum_{rs}^{3}M_{rs}\dot{Q}_r\dot{Q}_s.$$

The potential energy of the condensers is

$$\mathcal{E}_C=\frac{Q_1^2}{2C_1}+\frac{Q_2^2}{2C_2}+\frac{Q_3^2}{2C_3}+\frac{Q_4^2}{2C_4}.$$

Eliminating $Q_4$ we get

$$\mathcal{E}_C=\frac{Q_1^2}{2}\left(\frac{1}{C_1}+\frac{1}{C_4}\right)+\frac{Q_2^2}{2}\left(\frac{1}{C_2}+\frac{1}{C_4}\right)+\frac{Q_3^2}{2}\left(\frac{1}{C_3}+\frac{1}{C_4}\right)+\frac{Q_1Q_3-Q_2Q_3-Q_1Q_2}{C_4}.$$

The potential energy of the sources is

$$\mathcal{E}_E=-E_1Q_1-E_2Q_2-E_3Q_3+E_4Q_4.$$

Eliminating $Q_4$ we have

$$\mathcal{E}_E=Q_1(E_4-E_1)-Q_2(E_2+E_4)+Q_3(E_4-E_3).$$

The electrical dissipative forces are found as follows. Suppose that each of the charges is given an infinitesimal change, subject to the conditions that $Q_4=Q_3+Q_1-Q_2$, $\delta Q_4=\delta Q_3+\delta Q_1-\delta Q_2$, and $\dot{Q}_4=\dot{Q}_3+\dot{Q}_1-\dot{Q}_2$. The change in energy of the system due to heat produced in the resistances by these small changes in charge is given by

$$\delta\mathcal{E}=-[R_1\dot{Q}_1\delta Q_1+R_2\dot{Q}_2\delta Q_2+R_3\dot{Q}_3\delta Q_3+R_4\dot{Q}_4\delta Q_4].$$

Eliminating $\delta Q_4$ and $\dot{Q}_4$ we get

$$\delta\mathcal{E}=-\{\delta Q_1[R_1\dot{Q}_1+R_4(\dot{Q}_1+\dot{Q}_3-\dot{Q}_2)]+\delta Q_2[R_2\dot{Q}_2-R_4(\dot{Q}_1+\dot{Q}_3-\dot{Q}_2)]+\delta Q_3[R_3\dot{Q}_3+R_4(\dot{Q}_1+\dot{Q}_3-\dot{Q}_2)]\}.$$

Therefore, for a displacement of $\delta Q_1$ only, we have

$$\delta\mathcal{E}_1=F_1\delta Q_1=-\delta Q_1[R_1\dot{Q}_1+R_4(\dot{Q}_1+\dot{Q}_3-\dot{Q}_2)].$$

Hence the generalized dissipative force corresponding to $Q_1$ is

$$F_1=-R_1\dot{Q}_1-R_4(\dot{Q}_1+\dot{Q}_3-\dot{Q}_2).$$

Likewise,

$$F_2=-R_2\dot{Q}_2+R_4(\dot{Q}_3+\dot{Q}_1-\dot{Q}_2)$$

$$F_3=-R_3\dot{Q}_3-R_4(\dot{Q}_3+\dot{Q}_1-\dot{Q}_2).$$

If the inductances are free to move and if, perhaps, certain condenser plates are allowed to move, then the circuit regarded as a mechanical system possesses kinetic and potential energy. As previously mentioned, this energy must be included in the Lagrangian function. In order to represent this energy we must, of course, introduce a set of coordinates $q_1, q_2, \cdots, q_m$, sufficient in number to represent completely the configuration of the system. (If one coil only were allowed to move without rotation in the plane of the paper, then only two coordinates, say $x$ and $y$, would be required. If, in addition to this, one condenser plate were allowed to move, three displacement coordinates would be required.)

A general expression for the kinetic energy of a mechanical system referred to fixed axes is

$$T_{\text{mech.}}=\tfrac{1}{2}\sum_{rs}^{m}a_{rs}\dot{q}_r\dot{q}_s,$$

where $\dot{q}$ represents a generalized velocity corresponding to the generalized coordinate $q$ and

where the coefficients $a_{rs}$, containing the masses and in some cases the moment of inertia of the parts, are functions of the displacement coordinates. The indices of the double sum each run from 1 to $m$, $m$ being the number of displacement coordinates necessary to represent the configuration of the system. Referring to the double pendulum, $m=2$, $q_1=\theta$, $q_2=\varphi$, $a_{11}=(m_1+m_2)r_1^2$, $a_{22}=m_2r_2^2$, and $a_{12}=a_{21}=m_2r_1r_2\cos(\varphi-\theta)$.

The potential energy of the system, treated as a mechanical system, will depend entirely on the particular circumstances and can only be written explicitly when these circumstances are given. For example, if a set of springs is holding the coils and condenser plates and if, perhaps, certain parts are allowed to move under gravity, the potential energy can be written in explicit form. If the circuit is on a horizontal table with no springs or other conservative forces acting on the parts, the potential energy is, of course, zero. We shall denote the potential energy of the circuit, treated as a mechanical system, as $V'$.

In like manner the dissipative forces corresponding to the various coordinates $q_1, q_2, \cdots, q_m$, can only be stated explicitly when definite conditions are known. The dissipative

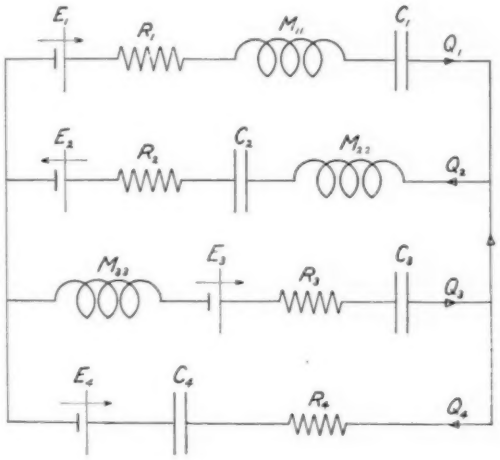



FIG. 6. Circuit to illustrate the more general form of Lagrange's equations. All condenser plates and all inductances may be allowed to move.

force acting on $m_1$ of the double pendulum is a typical example. Generalized dissipative forces corresponding to the displacement coordinates will be denoted by $f_1, f_2, \cdots, f_m$. In any actual problem any one or all of these may be zero.

We are now in a position to write down the Lagrangian function for the circuit. Including both electrical and mechanical energies, it is clear that

$$L=\tfrac{1}{2}\sum_{rs}^{3}M_{rs}\dot{Q}_r\dot{Q}_s+\tfrac{1}{2}\sum_{rs}^{m}a_{rs}\dot{q}_r\dot{q}_s-\mathcal{E}_C-\mathcal{E}_E-V'.$$

The Lagrangian equations for the system are, therefore,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{Q}_r}\right)-\frac{\partial L}{\partial Q_r}=F_r, \tag{3}$$

$$\frac{d}{dt}\left(\frac{dL}{\partial\dot{q}_r}\right)-\frac{\partial L}{\partial q_r}=f_r, \tag{4}$$

where, for the particular circuit in hand, $r$ in (3) extends from 1 to 3 and in (4), from 1 to $m$. When the operations of (3) and (4) are carried out there result $3+m$ second order differential equations. A solution of these simultaneous equations gives the complete electrical behavior as well as mechanical motion of the circuit. It is important to note that when inductance coils and condenser plates are allowed to move the expressions for mutual inductance and capacity become functions of the coordinates $q_1, q_2, \cdots, q_m$. This fact must be taken into consideration when performing the operations of (3) and (4).

In the more general case where a circuit is composed of any number of inductances, resistances, and condensers, any one or all of which are allowed to move, it is seen that the appropriate Lagrangian function is

$$L=\tfrac{1}{2}\sum_{rs}^{n}M_{rs}\dot{Q}_r\dot{Q}_s+\tfrac{1}{2}\sum_{rs}^{m}a_{rs}\dot{q}_r\dot{q}_s+\sum EQ-\tfrac{1}{2}\sum\frac{Q^2}{C}-V',$$

where it is assumed that, before $L$ is made use of (see the previous example), the number of charges have been reduced to a necessary minimum by Kirchhoff's law and also the displacement coordinates have likewise been reduced to the necessary minimum by the constraints on the system.

It is clear from the first two examples given, that the Lagrangian method of arriving at the differential equations of a simple stationary

electrical circuit offers no particular advantage over the usual Kirchhoff's law method. However, for a combination electrical-mechanical system, particularly where it may be advantageous to use coordinates for the moving parts other than the simple rectangular type (Lagrange's equations apply to *any system* of coordinates), the Lagrangian method reduces what otherwise might be very difficult to a more or less routine task. When the Lagrangian function $L$ has been properly determined, terms in the final differential equations are automatically given the proper algebraic signs. Likewise all interrelations of electrical and mechanical quantities are automatically taken care of.

Since electrical engineering and its related branches are replete with combination electrical-mechanical problems, it is believed that this method has a wide field of usefulness in modern engineering.

# Here and There

### New Appointments and Promotions

Dr. H. R. Crane has been promoted to an assistant professorship in the department of physics at the University of Michigan.

★

Professor John A. Wheeler of the University of North Carolina has been appointed assistant professor of physics at Princeton University.

★

Dr. Millard F. Manning of the University of Toledo has accepted an appointment as assistant professor of physics at the University of Pittsburgh. Professor Manning will teach primarily courses in the structure of metals.

★

At the Massachusetts Institute of Technology professors Robley D. Evans and Joseph C. Boyce have been promoted from assistant professors to associate professors; Drs. Walter Albertson, L. C. Van Atta, and C. M. Van Atta have been promoted from research associates to assistant professors; Dr. G. G. Harvey of the College of the City of New York has been appointed assistant professor; and Mr. J. H. Buck of the University of Rochester has been appointed instructor.

★

L. P. Delsasso has been granted a leave of absence for one year to serve as exhibits expert in the physical sciences for the University of California exhibit in the Golden Gate International Exposition.

Professor Maurice Ewing has been awarded a Guggenheim fellowship and will be on leave of absence from Lehigh University during 1938–1939.

★

The Burgess Battery Company has elected Mr. Dan W. Hirtle president and Dr. C. F. Burgess chairman of the board of directors at its annual meeting held March 26, 1938. Prior to his association with Burgess, Mr. Hirtle was an assistant superintendent of the American Rolling Mills Company, Middletown, Ohio.

★

Dr. Lee A. DuBridge, Harris professor of physics and chairman of the department, has been appointed dean of the faculty of arts and sciences of the University of Rochester. He succeeds Dr. Leonard Carmichael, who has been elected president of Tufts College.

★

### Summer Plans

Professor I. I. Rabi of Columbia University will be a member of the summer quarter staff at Stanford University.

★

### Sigma Xi Lecture

On Saturday, April 2, Dr. W. F. G. Swann, director of the Bartol Research Foundation, gave the annual Sigma Xi lecture at Carnegie Institute of Technology.

## Research Fellowship Appointments Announced

The first five Westinghouse Research Fellows have been selected for work in fundamental science at the Westinghouse Research Laboratories in East Pittsburgh. The successful candidates chosen from a group of fifty are:

*Robert O. Haxby*, University of Minnesota. Mr. Haxby, whose home is in Minneapolis, is twenty-six, has received the bachelor's and master's degrees in electrical engineering from the University of Minnesota. He has been a graduate student in physics there for the last four years and has participated actively in the nuclear physics research there. He expects to receive his doctor's degree in physics next June. He will join the group who will carry on research in nuclear physics, using the large high voltage generator now being built in East Pittsburgh.

*John A. Hipple*, Princeton University. Mr. Hipple is twenty-seven. His home is in Lancaster, Pennsylvania, where he received a bachelor's degree in physics from Franklin and Marshall College. He has been a graduate student in physics at Princeton during the past four years, where he has worked with Professor Walker Bleakney and where he expects to receive the degree of doctor of philosophy in June. He will design and construct a large mass spectrograph of high resolving power for studies in molecular structure in the Westinghouse Laboratories.

*W. E. Stephens*, California Institute of Technology. Stephens is twenty-six, and expects to receive the doctor's degree next June for researches in nuclear physics conducted in the Kellogg Radiation Laboratory in Pasadena. While in Pasadena, he has collaborated with Professor C. C. Lauritsen. He is a native of St. Louis, Missouri, where he received a bachelor's degree in 1932 and a master's degree in 1934 from Washington University. At Washington University he did researches on focusing of ion and electron beams in collaboration with Professor A. L. Hughes. In Pasadena, Stephens has studied the production of neutrons by bombarding boron and nitrogen with high energy deuterons. He has also worked on the development of a new type of high voltage discharge tube. In East Pittsburgh, he will continue his researches on nuclear physics with the large generator at the Westinghouse Laboratories.

*Sidney Siegel*, Columbia University. Dr. Siegel is twenty-six and received his doctor's degree two years ago for researches on elastic and magnetic properties of metals carried out under the direction of Professor S. L. Quimby. Dr. Siegel is a native of New York City, and received his bachelor's degree in 1932 from Columbia College. Dr. Siegel will continue his studies on fundamental properties of metals. He plans to develop methods of growing single crystals of pure metals, and of alloys which show order-disorder transitions in order to study how the elastic properties of alloys depend on the arrangement of the atoms in the crystal lattice.

*W. E. Shoupp*, University of Illinois. Dr. Shoupp is at present an instructor in physics at the University of Illinois, where he is actively engaged in studies on nuclear physics, with the Illinois cyclotron. Dr. Shoupp is thirty, and his home is in Troy, Ohio. He received his bachelor's degree from Miami University, Oxford, Ohio, and went to Illinois for graduate study in 1931. In addition to his work in nuclear physics, he has made several contributions to extreme ultraviolet spectroscopy as a collaborator of Professor P. G. Kruger. At the Westinghouse laboratories, he will continue studies in nuclear physics, working with Dr. W. H. Wells on the high voltage generator.

The five appointments are the first to be made under a plan announced last December, whereby the Westinghouse

Electric and Manufacturing Company plans to support the work of ten young physicists on fundamental studies broadly related to the electrical industry. The fellows are appointed for one year and are eligible for one reappointment. They will work under the general supervision of Dr. E. U. Condon, Associate Director of the Laboratories.

"We have been extremely gratified by the response to the fellowship plan which is represented by the splendid group of applications received," said Dr. L. W. Chubb, Director of the Research Laboratories, in discussing the appointments of the first group of fellows. "This made the task of selecting the men a difficult one. We regret only that it is not possible to offer appointments to a larger number of the applicants."

★

### The Electrochemical Society

The Electrochemical Society held its Spring meeting at Savannah, Georgia, April 27 to April 30, 1938. The Society met jointly with paper experts to review and discuss old and new problems with which that industry is confronted. The symposia were scheduled on industrial applications of electro-osmotic phenomena and on the electro-deposition of metals, particularly manganese, chromium, zinc, copper, nickel and alloys of copper and of nickel. A fourth session was devoted to principles of electrochemistry including papers on the dropping mercury electrode, *p*H determination, new differential dynamometer, motor-electrolytic potentials, film potentials, and over-voltage.

★

### Institute of Radio Engineers

The Institute of Radio Engineers cordially invites all interested physicists and engineers to attend its Annual Convention at the Hotel Pennsylvania, New York, June 16th–18th. At the technical sessions, fifty papers are scheduled for presentation and discussion on vacuum tubes, communication circuits, television, and other subjects of value to workers outside the radio field.

In addition, a comprehensive display of laboratory instruments, communication-engineering equipment, vacuum tubes, and materials will be open for inspection throughout the convention. Thirty leading manufacturers are participating in this exhibition.

The June issue of the *Proceedings of the I. R. E.* will carry the complete program with summaries of the technical papers or a list of papers to be presented may be obtained from the I. R. E. Convention Secretary. There will be no registration fee or other charge.

★

The Institute of Metals, London, has made the first award of its new medal to Sir. William Bragg, president of the Royal Society. The presentation was made on March 8. The medal—in platinum—is given for outstanding services to nonferrous metallurgy.

*Science*, March 18, 1938

### New Research Division

The formation of a division of engineering and research by Crane Co., Chicago, with L. W. Wallace, head of engineering research for the Association of American Railroads, as director has been announced by Charles B. Nolte, president of Crane Co. The new Crane division will comprise the existing division of research and development and the product engineering department of the company.

"The new division has been formed to coordinate all engineering activities of the company," said Mr. Nolte, "and to further its progress in diversified fields. The importance and growth of this technical work have proved the advisability of maintaining a complete and well-staffed engineering division, free from all responsibilities except those relating to engineering, research, design and experimental work."

★

### Engineering Lectures

Dr. Reinhold Rudenberg, formerly chief electrical engineer of the Siemens-Schuckert Company and honorary professor at the Technische Hochschule in Charlottenburg, now with the General Electric Company of England, gave a series of three lectures at Harvard University on the following dates:

May 11, **Technique and economy of electric high power transmission over long distances.**
May 12, **Oscillations common to synchronous generators and asynchronous motors in disturbed networks.**
May 13, **Straining of machine windings by short circuit forces and traveling waves.**

★

### Industrial Research Institute

A new Industrial Research Institute is being organized by the Division of Engineering and Industrial Research of the National Research Council. The purpose of this institute is to consider general problems which will be of interest to a large number of research laboratories such as the selection, training, handling, and incentive and bonus plans for research personnel. They plan also to consider the problem of securing adequately trained industrial research workers, the organization, administration and operation of budgets of the research laboratories in relation to that of the other departments of a company, and the coordination of research activities with market studies and sales promotion. Recently an executive committee was elected with Robert B. Colgate as chairman. Further information concerning this institute may be obtained from Maurice Holland, who is serving as executive officer.

★

### Necrology

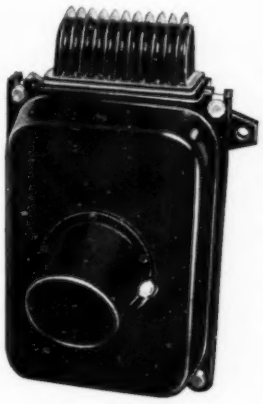
We regret to announce that Professor P. A. Ross of Stanford University died March 13, 1938.

# Innovations in Instruments

### RCA Television Parts

RCA has recently made available certain specialized television parts for use by experimenters within radius of television transmitting stations. With these new parts and other standard parts already available, it is now possible for the amateur experimenter who is equipped with sufficient technical knowledge, to assemble his own Kinescope deflecting circuits for use in experimental television receivers. The new parts listed for sale include a deflecting yoke, two power transformers, a vertical output reactor, a vertical oscillation transformer, a horizontal oscillation transformer, a horizontal output transformer, two power supply capacitors, and a power supply reactor. A folder describing these parts and giving wiring diagrams is available from RCA Parts distributors or from the RCA Parts Division, Camden, New Jersey.

★



### Photo-Tube Relay

Built for outdoor service in all kinds of weather, this is a new photo-tube relay unit developed by G-M Laboratories, Chicago. Complete protection from rain and snow is given by the heavy watertight case and lens visor. A specially designed resistor is mounted in the top of the case, with fins for dissipating heat. The unit is designed for 115 volt d.c. operation.

★

### d'Arsonval Motion Picture

The French producer, Jean Benoit-Levy, has released a moving picture on d'Arsonval's discoveries. Although d'Arsonval is over eighty-four years of age, he plays his own title role. Professor Barail presented the picture to the public on Wednesday, April 20, 1938, at the McMillin Academic Theater of Columbia University. The performance was given under the auspices of the Institute of Arts and Sciences and the Institut des Études Françaises of Columbia University.

★

### Solving Machine-Tool's Special Problems

The Westinghouse Technical Press Service announces a squirrel-cage induction motor with an adjustable no-lead speed. Ordinary induction motors come quickly to almost synchronous speed, and voltage control or high rotor resistance do not appreciably affect the light-load speed. A novel motor and control has been devised which accomplishes the very unusual result of permitting an induction motor to be run at light-load at any desired low speed. The scheme consisted essentially of internally loading the motor. The motor has a three-phase delta winding, with one phase disconnected from the other two, making it essentially an open-delta motor. When some lower light-load speed is desired the third phase is connected to the line in reverse. The current through it is in a generating instead of motoring direction and tends to buck the motor action of the two open-delta phases. The motor slows down as though a load had been applied to motor with a high resistance rotor. By controlling with a rheostat the current through the reversed phase the bucking action and, thereupon, the motor speed can be closely controlled. This motor has been devised especially for chucking work into a large lathe where it is necessary to close in the chuck jaws firmly but without shock.

★

### New Rectichargers

The Raytheon Manufacturing Company announces a new automatic charging device which does everything necessary to keep a storage battery fully charged except

adding the water to make up for evaporation. The control unit is designed to adjust the charging current rate in proportion to the current drawn from the battery. Overcharging of the battery is prevented by this unit and the battery will be kept fully charged. The rectifier unit in this charger is a copper oxide one. The complete charger is mounted in a metal cabinet.

### Artificial Lightning at World's Fair

General Electric's high voltage laboratory at Pittsfield, Massachusetts, where man-made lightning has been demonstrated will be transported to the New York World's Fair in 1939. The general public will have an opportunity to see in action this artificial phenomenon, created for the purpose of duplicating the forces of natural lightning. The laboratory will be housed in Steinmetz Hall, named after the late Dr. Charles P. Steinmetz, credited with being the father of man-made lightning. Present plans call for the display of a 10,000,000-volt bolt of lightning leaping across a gap of 30 feet with a crash greatly simulating the thunder which accompanies natural lightning. There will also be the spectacular million-volt, three phase arc, whose ribbons of electrical fire will twist their way for 30 feet in the air above the apparatus. There will also be the unusual whirling arc in which a million volts jump between three points with vivid flashes also leaping high in the air.

★

### New Substage Lamp

The Spencer Lens Company has recently developed a new microscope substage lamp which has a Bakelite molded housing. This housing is so strong that it may be knocked off the table and stood upon without breaking.

# Some of the New Books

**Kinetic Theory of Gases.** EARLE H. KENNARD, professor of physics, Cornell University. First Edition. The McGraw-Hill Book Company, New York, 1938. xiii+483 pp. 94 figs. Price, $5.00.

The author remarks in his preface that although the kinetic theory of gases is a branch of physics which has passed from the stage of excitement and novelty into staid maturity, there remains a demand for adequate exposition of the subject. To meet this demand he has provided a clear and concise presentation of kinetic theory, from which the illogical and unsystematic elements have been excised so far as possible. While primarily classical in point of view, the book treats in detail certain reinterpretations of kinetic theory demanded by quantum mechanics. These reinterpretations are of particular significance in connection with collision processes, specific heats, and the properties of the ideal gas from the standpoint of the Bose-Einstein and Fermi Dirac statistics.

After an introductory chapter devoted to definitions and fundamental concepts, the general distribution laws are derived and discussed by conventional methods. In a chapter on the spatial distribution of molecules there is a thorough treatment of molecular scattering from the wave mechanical as well as from the classical point of view. The middle portion of the book is chiefly concerned with the theory of transport processes, the equation of state, and the thermodynamic functions of real and ideal gases. An interesting chapter on statistical mechanics, in which its foundations are critically discussed, is followed by a detailed account of the wave mechanics of gases. The book concludes with a chapter on the electric and magnetic properties of gases.

The author is to be commended for his competent and thorough treatment of a field, which, though not new, still presents many interesting problems. The book should prove particularly useful as a text in kinetic theory for seniors and graduate students engaged in the study of physics and chemistry.

JOHN G. KIRKWOOD
*University of Chicago*

**Eisenlose Drosselspulen (Air Core Coils).** J. HAK· Pp. 316+viii. Figs. 253, Tables 32, $16\frac{1}{2}\times23\frac{1}{2}$ cm. K. F. Koehler Verlag, Leipzig, 1937.

It has long been a source of annoyance in the design of air core coils for various purposes to find the pertaining information widely scattered over many periodicals. The above book not only represents a very intelligent collection of practically all the formulas developed since Maxwell's famous treatise on *Electricity and Magnetism*, but it also gives a critical survey of their utility emphasized by many numerical tables for the useful ranges. All the formulas are carried back to their origin and written in a consistent uniform style, are illustrated by 65 examples chosen from all fields of application and presented frequently in graph form. As a novel feature, all the figures are collected at the end of the text and furnished with English captions. In addition, a German-English glossary of the more important terms adds to the usefulness for readers not quite familiar with the German technical expressions.

The arrangement of the text is logical, starting with the computation of self-inductance of circular and polygonal coils based upon the Neumann formula. The influence of insulation is briefly covered, as well as the important aspect of the most economical design. This is followed by the more complex computation of mutual inductance betweens coils of circular, rectangular, polygonal cross sections and various relative positions. Mechanical forces, effect of heating, skin effect and distributed capacity are well taken into account in separate chapters and many special coils are investigated, as for example, the spherical coil, coils of stranded wire, ring coils, and others; examples of construction are illustrated by pictures. Finally, tests on coils are briefly summarized and in an appendix the most recent form of the dust core coils is described. The thoroughness of this monograph is best illustrated by the extensive bibliography containing 687 references covering the literature up to early 1937. The book should prove most useful,

indeed indispensable, for all those who have to design and use coils for common as well as special purposes.

ERNST WEBER
*Polytechnic Institute of Brooklyn*

### Booklets Recently Received

The Ohio Chemical Manufacturing Company of Cleveland, Ohio, has issued a 16 page booklet dealing with equipment for control and storage of high pressure gases.

**Metallographic Equipment and Accessories, Catalog E-225,** published by Bausch and Lomb Optical Company of Rochester, New York. This 32 page booklet contains descriptions of Ils Metallographic Equipment, Si Metallographic Equipment, and Vertical Illuminator.

**The General Radio Experimenter.** (8 pp.) Vol. XII, No. 10, March, 1938. This issue contains articles on "Modern Improvements for the Beat-Frequency Oscillator," "Variacs Used as Controls in Flexible Stage-Lighting Unit," and "Measurng 100,000 Rpm."

**Oscillographer.** (8 pp.) Vol. 2, Nos. 1–2, March–April, 1938. In this issue are articles on "Type 169 Nine Inch Laboratory Oscillograph," and "The DuMont Simplified Television System Demonstrated."

# Contributors to This Issue

**Dr.** William T. Richards received his A.B. degree from Harvard in 1921 and his Ph.D. from there in 1924. He was assistant to the director of Gibbs Memorial Laboratory from 1923 to 1924. He held an International Education Board fellowship from 1924 to 1926. He was an instructor at Princeton from 1926 to 1928 and later was appointed assistant professor. Dr. Richards is now with the Rockefeller Institute for Medical Research in New York.

**Dr.** Ernest G. Linder received his B.A. degree in 1925, and his M.S. in 1927 from the State University of Iowa. In 1931 he received his Ph.D. from Cornell University. He was an instructor in physics at the University of Iowa from 1925 to 1927, and at California Institute of Technology from 1927 to 1928. From 1928 to 1932 he was a Detroit Edison Research Associate at Cornell. Since 1932 he has had the position of research physicist, Research Division, RCA Manufacturing Co., Inc.

**Dr.** Hans Klemperer received his doctor's degree in electrical engineering from the Institute of Technology in Berlin in 1929. He was an instructor at the University of Aachen until 1930 when he accepted a position in the research laboratories of Siemens and Halske in Berlin. In 1933 he worked in the development department of the Siemens Schuckert Werke in Berlin, in charge of applications of grid controlled rectifiers. In 1934 he was connected with the Radio Corporation in the electronic tube laboratory at Harrison, New Jersey. Since 1935 he has been with the Research Laboratory of Westinghouse Electric and Manufacturing Company at East Pittsburgh, Pennsylvania, working on ignitron and gaseous tube development.

**Mr.** R. Weller received a B.S. from M.I.T. in 1927 and subsequently taught mechanical subjects at Mechanics Institute in Rochester, New York. He obtained an M.S. in physics from Ohio State University in 1934 and has been engaged in photoelastic research in the department of physics and in the Engineering Experiment Station of Ohio State for several years.

**Dr.** George H. Shortley received his bachelor's degree in electrical engineering at the University of Minnesota in 1930 and his Ph.D. in physics from Princeton in 1933. He was subsequently a National Research Fellow at M.I.T. and Harvard and is now assistant professor of physics at Ohio State University. His principal research has been in the field of the theory of atomic spectra and he is co-author of the book, *Theory of Atomic Spectra*, by Condon and Shortley.

# Contributed Original Research

## Contribution to the Problem of Arcback in High Voltage Discharge Devices

H. KLEMPERER
*Westinghouse Electric and Manufacturing Company, East Pittsburgh, Pennsylvania*
(Received February 4, 1938)

It has been found in rectifier and ignitron practice that arcback can occur after deionization of the discharge path at such vapor pressures where no self-sustaining glow can develop. Experiments to study this kind of an arc breakdown were made in Hg vapor vessels between cold electrodes at voltages between 10 and 60 kv. It appeared that the start of arc breakdown was not dependent on the Paschen law. At the voltages used and distances between the electrodes of less than 50 cm arc breakdown always started at PD values which were much lower than those required to sustain a glow discharge. This is contrary to earlier experience at lower voltages. With rising voltage arc breakdown frequency increased in an exponential manner. Changes in vapor pressure and electrode spacing had no definite effect on arc breakdown frequency as long as glow discharge conditions were not approached. Changes in partial air pressure did not affect arc breakdown frequency if air pressure was below 10 percent of mercury vapor pressure. Heat treatment of the cathode (graphite "anode") has a large effect in reducing arc breakdown. Even well heat-treated "anodes" show seasoning effect, i.e. steady decrease of arc breakdown frequency after initial maximum. Phenomena on cathode surface are believed to be responsible for formation of arc breakdown.

### DESCRIPTION OF PROBLEM AND EARLIER EXPERIENCE

THE phenomenon which is commonly called "arcback" and which will be called "arc breakdown" in some of the following discussions, is noticed as a sudden appearance of an arc spot on the "anode"[1] of a rectifier during the inverse part of the cycle, i.e. when the "anode" is negative. The discharge which is observed following breakdown shows the typical properties of the arc discharge: low arc drop and high current density in front of the cathode.

No direct experiments showing onset of arc breakdown have been made so far. Cathode-ray oscillograms[2] which, due to inertia of the triggering relay (Norinder type) started after elapse of 3 microseconds, did not reveal any sign of a preceding phase of the discharge, during which voltage drop was higher.[3] The highest speed type of the cathode-ray oscillograph has not yet been used for such investigation, as far as the author knows, and it may be expected that faster oscillograms would provide insight into the formation period of the arcback discharge.

Though the problem of arcback is as old as the mercury arc rectifier itself, no complete theory has been formed so far and rectifier and tube designing to a large degree is still empirical. In early days poor vacuum, which caused the development of a glow discharge[4] during the inverse part of the cycle, was held responsible for the formation of arcback. The vacuum was improved with the development of better pumps and vacuum technique, and the number of arcbacks decreased. But they did not disappear. Later on ion diffusion to the "anode" was made responsible for arcback[5] and shielding of the anodes and deionizing grids greatly improved the quality of the rectifier. But backfires

[1] "Anode" and "cathode" in quotation marks mean mechanical parts following rectifier practice, and do not denote the electrical polarity.

[2] Oscillographic investigation by Maxfield, published by Slepian and Ludwig, Trans. A. I. E. E. **51**, 92 (1932). For oscillograph see O. Ackermann, Trans. A. I. E. E. **49**, 467 (1930).

[3] See Rogowski, Zeits. f. Physik **100**, 1 (1936).

[4] R. Seeliger, Naturwiss. **16**, 655 (1928).

[5] J. v. Issendorff, M. Schenkel and R. Seeliger, Wiss. Veröff. a.d. Siemens Werken, **9**, 73 (1930).

were not eliminated. Then the ignitron was believed to be the ideal solution of the backfire problem,[6] since it does not use "keep alive" means during the inverse part of the cycle and ion diffusion from other arcs is eliminated. But even with the ignitron arcbacks were observed, and these arcbacks occur not only during the transition period, but also near the time of the inverse voltage peak, when residual ionization practically has disappeared.

As mentioned, some arcbacks—mostly those in high current rectifiers—happen during the transition periods.[7] These arcbacks apparently are associated with residual ionization of the discharge path, which was conducting a short time previously. Other arcbacks, particularly those encountered in high voltage rectifier and tube work, occur near the time of the inverse voltage peak with very little or no ionization present. It is difficult to connect this type of arcback to any such phenomenon[8] as known to start a self-sustaining discharge. Ion bombardment of the cathode is very small, thermionics are hardly involved in the initial phase, and the field strength at the cathode surface is of the order of 1000 to 10,000 v/cm, which is very small compared to gradients at which cold emission is to be expected.

## Range of Testing, Set-Up and Procedure

Breakdown between cold electrodes, with no voluntary source of ionization present, and moreover under voltage, vapor pressure and spacing conditions where neither the development of a (Townsend) glow nor cold emission is expected may have some other relation to field strength. With this in mind the author carried out a number of breakdown tests, applying high unidirectional voltage to electrodes of test vessels. These vessels contained a graphite "anode," a mercury pool "cathode" and no voluntary source of ionization. Voltages between 15 and 60 kv were used. The Hg vapor pressure was varied between 0.26$\mu$ (5°C) and 89$\mu$ (80°C) and electrode spacings ranged from 1 cm to 50 cm (sometimes 100 cm).

The circuit diagram of the test set-up is given in Fig. 1. Engineering experience seems to indicate that high voltage systems with a limited power supply are less liable to show breakdowns than high power systems of same voltage and electrode spacing.[9] The probable reason is that a low power system does not supply current sufficiently rapidly for the arc spot to develop after the breakdown is initiated, and that with some unnoticed glow or flash the failure may clear itself. Similar reasoning applies to low pressure breakdown, if the breakdown is detected by the following arc discharge current.[10]

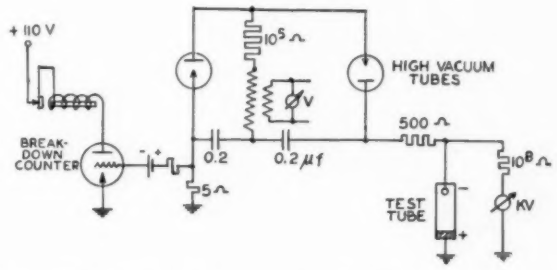



FIG. 1. High voltage set. (0.2 should read 0.2$\mu$f; kv should be meter.)

The current density in the cathode spot of an arc discharge lies around 5000 amperes/cm$^2$. The cathode spot of the arc, formed by breakdown as seen from visible marks on the "anode," sometimes is as small as $\frac{1}{5}$ mm in diameter. Therefore, the high voltage test set should be able to supply a current of 5 to 10 amperes during the first few microseconds. As shown by sketch the test set used a condenser discharge to feed the breakdown. Time constant was $T=50$ microsecond; initial current at lowest voltage (15 kv) was 30 amperes.

A sketch of the test vessels used is given in Fig. 2. Vessel *A* and *B* were water-cooled steel tanks with fixed electrode distances. Vessel *C*

[6] W. G. Dow, *Fundamentals of Engineering Electronics* (J. Wiley 1937), p. 488. J. Slepian and L. R. Ludwig, "A New Method for Initiating the Cathode of an Arc," Trans. A. I. E. E. **51**, 693 (1933).

[7] Ludwig, Maxfield and Toepfer, Trans. A. I. E. E. **53**, 75 (1934).

[8] Sometimes radiation may have some influence; there is no reason against the consideration of cosmic rays as a cause of some of the breakdowns. For other mechanism see reference 17.

[9] Compare spacing in a 100 kv x-ray set with spacing used for power stations of same voltage.

[10] In high voltage rectifier and tube work the number of arcbacks observed depends on sensitivity of back current measuring device.

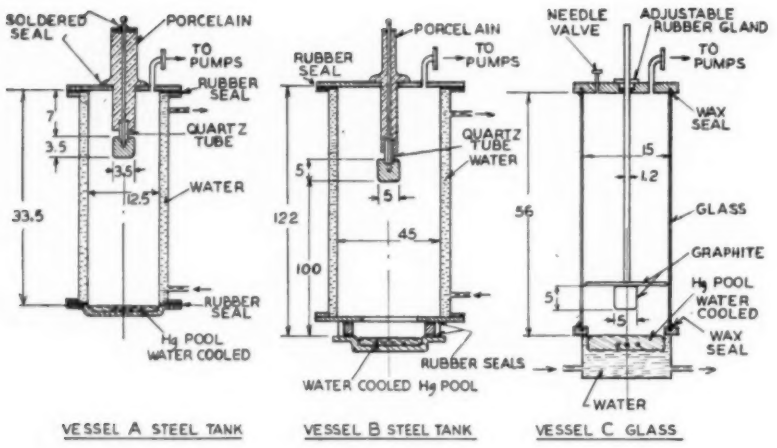



FIG. 2. Test vessels. Dimensions in cm.

had glass walls and the electrode spacing was variable between 1 and 50 cm.

Testing was done in the following way: the vessel was evacuated to a certain residual gas pressure which was read with the McLeod gauge. The temperature of the vessel was held constant at a desired point by regulating the cooling water or using a furnace. When conditions seemed to be constant direct voltage was applied in the direction prevailing during arcback periods ("anode" negative) and the occurrence of breakdown was watched or counted over a certain length of time. Records of the breakdown number were made by an electronic counter. These tests were repeated for a great number of times in the hope of obtaining reproducible average results.

## DISCUSSION OF RESULTS

### (a) PD law not applying to breakdown

Right from the beginning of the tests it appeared that the pressure distance relation (Paschen law or PD law), which states that a discharge starts at a constant voltage if the product of gas pressure × electrode spacing (PD) is kept constant, does not apply to arc breakdown. With fair accuracy a glow discharge could repeatedly be started at the same voltage if pressure and distance were changed as the law prescribes. The start of arc breakdown, however, did not follow any such relation.

Figure 3 represents this situation in the voltage pressure distance (VPD) diagram. The glow curve[11] on the left side of the minimum voltage (around 600 v for Hg vapor) is only reproducible for the lowest values of voltage.[12] In every case at voltages above 15 kv and at spacings below 50 cm breakdown occurred at PD values where no glow discharge could be observed. This is in accordance with above mentioned experience on high voltage ignitrons, were arcback was observed frequently near the peak of inverse voltage, and very seldom during the transition period. On the other hand, at lower voltage, smaller spacing ($D=4$ cm) and accordingly higher vapor pressures Marshall and Toepfer[13] observed glow discharges before any breakdown started. Some of their results are given in the lower part of the VPD diagram of Fig. 3. This situation seems to apply to some arcback phenomena in regular power rectifiers, which run at large current and moderate voltage (600 v for instance). Here, in many cases, arcback was notably preceded by a glow discharge.

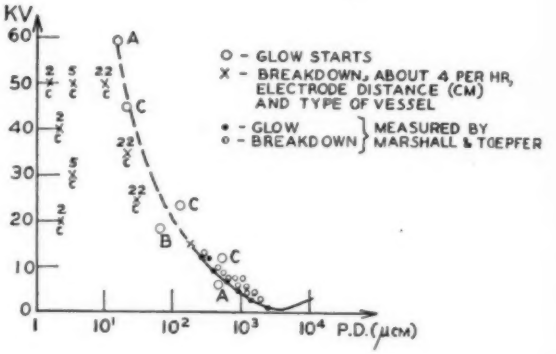



FIG. 3.

[11] It is more exact to say the "$\mu=1$ curve." $\mu$ is the rise in ionization, defined $\mu=\gamma(\epsilon^{aL}-1)$. See Rogowski, Zeits. f. Physik **100**, 15 (1936).

[12] Smede, Electric J. **25**, 403 (1928).

[13] Marshall and Toepfer, test results, partially published by Slepian and Ludwig, Elect. Engineering **50**, 793 (1931).

### (b) Influence of voltage on breakdown frequency

The well-known increase in breakdown frequency with rising voltage appeared from all experiments without exception. It was not possible to find a general mathematical relation because in most of the tests the probability of arc breakdown changed during the tests, even after hours of waiting. Especially if the voltage was high it was difficult to get repeatable average values. In general, the rise of backfire frequency with voltage had an exponential character,[14] and in one case where a commercial high voltage rectifier was subjected to the described direct voltage tests, results checked fairly well with the law, as given by Fig. 4. In this test all "anodes" of the rectifier were tied together to test voltage; all grids were kept floating and "cathode" and tank were tied to ground. The rectifier was

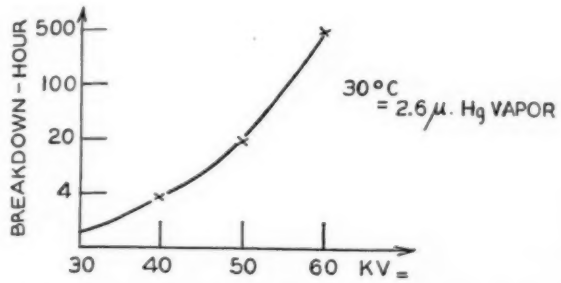



FIG. 4. Anode voltage *vs.* breakdown frequency of commercial mercury arc rectifier.

thoroughly treated and aged, and the above readings were taken after several hours of testing.

### (c) Influence of vapor pressure on breakdown frequency

No definite tendency of the breakdown frequency to increase as a result of rising vapor pressure could be observed. If conditions were approached, under which a self-supporting glow discharge can exist, the breakdown frequency rose rapidly. But in the region below this happening the backfire frequency drifted erratically up and down, and other uncontrolled influences overshadowed any effect of vapor pressure.[15]

Fig. 5 gives results of such measurements taken on vessel *A*. Similar observations were made with all the other test vessels.

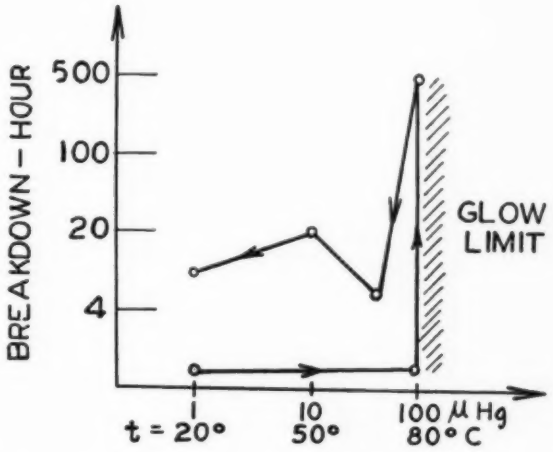



FIG. 5. Breakdown frequency *vs.* vapor pressure at 60 kv $d = 22$ cm (taken with vessel "A" on Nov. 27, 1935).

### (d) Influence of electrode spacing on breakdown frequency

Similar erratic relations were observed between electrode spacing and the breakdown frequency. If the geometrical configuration leaves unobstructed a path, along which a glow discharge can start, the backfire frequency increases. If such conditions exist, shields, baffles or grids which suppress the start of a glow from the upper side of the "anode" or from an unshielded "anode" stem also reduce the frequency of breakdown. On the other hand, if the electrode spacing becomes small enough to raise the field strength up to cold emission limits, breakdown follows. This effect has been observed with rough, untreated "anodes" at distances around 1 cm and voltages as low as 50 kv. But between these two borders, glow and cold emission, no definite tendency of the breakdown frequency in relation to spacing could be observed. Fig. 6 gives a sketch of test vessel *C*. A metallic shield was placed inside the upper part of the glass cylinder and tied to "anode," stem and cover plate. Lines of force are given on one side, breakdown

[14] Slepian and Ludwig, Trans. A. I. E. E. **51**, 92 (1932).
[15] This experience should not be applied to general rectifier practice without considering other effects of increased vapor pressure: the increase in deionization time, the change of arc drop, and maybe increase in leakage rate.

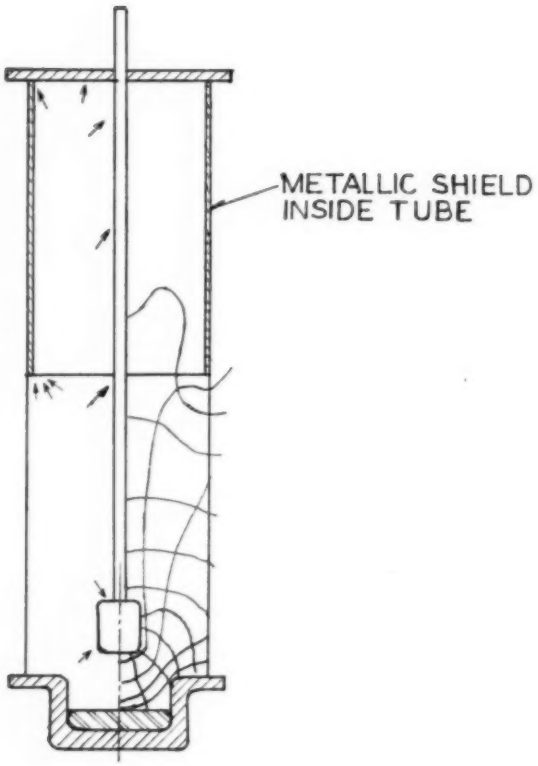



FIG. 6. Breakdown frequency and lines of force in vessel $C$ with metallic shield. Arrows indicate spots at which backfires occurred. Tests made at 50 kv, 2$\mu$ air, 30°C Hg.

spots, as observed by experiment, are indicated on the other side. Apparently some of the breakdown started from points with higher field strength, (edges), others developed along the line of the longest path. Repeated tests gave the same general situation but, of course, not exactly the same distribution of breakdown spots.

### (e) Influence of the partial air pressure on breakdown frequency

A great number of tests was run to find out the influence of air admixtures to mercury vapor on the breakdown frequency. Air was made leaking in from the atmosphere with a needle valve, and in other cases the slow development of a leakage pressure was waited for and measured with the McLeod gauge. The results were equally indefinite and hard to repeat. If the percentage of foreign gas was less than 10 percent of the mercury pressure no effect on breakdown frequency was observed. If the foreign gas pressure was higher than 10 percent the breakdown frequency was sensitive to any change of the foreign gas pressure. Mostly the breakdown became more frequent with increase in admixture, but there were cases observed when the effect was reversed. The minimum breakdown voltage of air is about one-half of that of mercury vapor, and the corresponding PD value (330$\mu$ cm) is ten times lower than that of Hg, so that, as far as glow is concerned, we should expect a considerable influence of small air admixtures. But for the breakdown frequency no such influence of partial air[16] pressure can be stated from the described tests.

### (f) Influence of heat treatment and aging on breakdown frequency

It is known from practice that a well-designed rectifier with good vacuum inside would show poor arcback qualities if untreated anodes were used. On the other hand, poorly designed rectifiers have been brought up to a fair performance simply by subjecting the anodes to a severe heat treatment and aging process.

Heat treatment consists in running the anodes up to 1400 or 1800°C in vacuum. Even after such treatments new anodes show initial backfires which become more and more seldom in frequency with the time of use, provided that no damage is done to anode surface or vacuum. Untreated anodes show this "seasoning" effect to a much larger degree and do not reach the final quality of heat treated anodes.

Both procedures, heat treatment and aging, evidently do the same thing: they remove backfire causes from the "anode" surfaces. Whether these causes consist of thin local dirt layers,[17] gas occlusions, irregularities in surface structure or loose particles, is still a matter of speculation.

---

[16] For the effects of small $H_2$ percentages see Dällenbach and Gerecke, E. T. Z. **57**, 937 (1936).

[17] A. Guentherschultze and H. Schnitger, Zeits. f. Physik, **102**, 34 (1936). Malter, Phys. Rev. **50**, 48 (1936). The possible relation of this type of low voltage-cold emission discharges across thin layers ("jet discharges") to the formation of arc breakdown will be discussed later.

# Effect of High Energy Electron Random Motion Upon the Shape of the Magnetron Cut-Off Curve

E. G. LINDER
*RCA Manufacturing Co., Inc., Camden, N. J.*
(Received March 2, 1938)

The shape of the cut-off curve is explained on the basis of a component of high energy random motion superimposed on the electron orbital motion (see references). Electron temperatures determined from the cut-off curve shape agree with those determined by probe methods. The possible effects of the random motion on other magnetron characteristics are pointed out.

A GREAT many experimental investigations have been made of the shape of the magnetron* cut-off curve, but no one appears to have attempted to determine what its shape should be from theoretical considerations. However, Hull's[1] work on the motion of electrons in magnetrons, and some experimental data on the effects of space charge, permit certain qualitative conclusions to be drawn immediately.

In a perfectly symmetrical cylindrical magnetron, neglecting space charge effects and the initial velocity of the electrons, and assuming that the electrons move only under the action of the electric and magnetic fields as described by Hull, the cut-off curve would have the general shape shown in Fig. 1(a). The anode current $I_p$ remains constant until the magnetic field reaches the critical value $H_c$ when $I_p$ drops abruptly to zero. This is evident from the equations of the electron path. If space charge is considered, $I_p$ no longer remains constant as $H$ becomes greater. The increasing curvature of the electron trajectories, and the longer transit time, cause an increase in space charge and a consequent decrease in $I_p$ before the cut-off point is reached. However, at $H=H_c$, $I_p$ drops abruptly to zero as before, as is shown in Fig. 1(b). If initial velocities also are considered the curve takes the form shown in Fig. 1(c). The first part of the curve, for $H<H_c$, is the same as in the previous case. But at $H=H_c$ an abrupt drop no longer occurs. At that point only those electrons having zero initial velocity fail to reach the anode, those having higher initial velocities require fields greater than $H_c$ for cut-off. This yields a curve of the form shown, which is the form always observed experimentally if the emission current is sufficiently large to cause appreciable space charge.

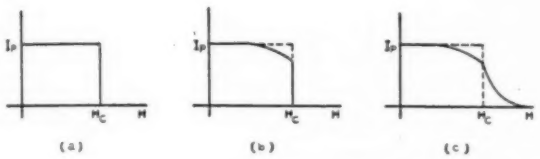



FIG. 1. Idealized magnetron cut-off curves. (a) With no space charge or initial velocity; (b) Space charge but no initial velocity; (c) With both space charge and initial velocity.

The first part of the cut-off curve, i.e., $H<H_c$, has been discussed previously by L. Tonks,[2] and more recently in greater detail by G. Grünberg and V. Wolkenstein,[3] and will not be dealt with in the present article which will discuss only the second part, i.e., $H>H_c$, the shape of which depends upon initial velocities.

Previous work[4, 5] has shown that electrons in magnetrons may acquire a Maxwell-Boltzmann distribution of velocities superimposed upon their orbital motion, and therefore behave very much as though emitted from a cathode of very high temperature. Probe measurements have shown that these temperatures are many times that of the actual cathode being of the order of 40,000°C for a plate potential of 100 volts in tubes investigated by the writer. Initial velocities corresponding to such temperatures should affect very

* The type of magnetron considered in this paper consists of a cylindrical anode having a concentric filamentary cathode, aligned axially in a uniform magnetic field.

[1] A. W. Hull, Phys. Rev. **18**, 31 (1921).

[2] L. Tonks, Physik. Zeits. Sowjetunion **8**, 572 (1936).

[3] G. Grünberg and V. Wolkenstein, Tech. Phys. USSR **4**, 179 (1937).

[4] I. Langmuir, Phys. Rev. **26**, 585 (1925).

[5] E. G. Linder, Proc. I. R. E. **26**, 346 (1938).

considerably the electron orbit and the shape of the cut-off curve.

An expression will be derived giving $I_p$ as a function of $H$ and the temperature $T$ of the electron swarm in the cathode neighborhood. In order to derive this expression, assume that all electrons start so near the central axis of the tube that only radial and axial components of initial velocity need be considered, components tangential to the cathode being negligible. Assume also that the initial velocities correspond to the temperature $T$ of the Maxwellian distribution.

The plate current due to electrons of number $dn$ having initial radial velocities between $u$ and $u+du$ is given by

$$dI_p = ev\,dn, \tag{1}$$

where $v$ is the radial velocity at the plate.

From Hull,[1] the radial velocity at any distance $r$ from the origin is given by

$$v = \frac{dr}{dt} = \left[\frac{2eV}{m} - \left(\frac{He}{2m}\right)^2 r^2 + u^2\right]^{\frac{1}{2}}, \tag{2}$$

where $V$ is the potential at any point, and $H$ is the magnetic field strength.

The number of electrons having initial radial velocity components between $u$ and $u+du$ according to Maxwell's distribution law, is given by

$$dn = 2n_0\epsilon^{-x^2}x\,dx, \tag{3}$$

where $x=(m/2kT)^{\frac{1}{2}}u$, and $n_0$ is the total number emitted. The axial components need not be considered since they do not contribute to the plate current.

The cut-off value $H_c$ of the magnetic field is found by putting $dr/dt=0$, for $r=r_p$, and $u=0$, in (2),

$$\text{i.e.}\quad H_c = (8mV_p/er_p^2)^{\frac{1}{2}}. \tag{4}$$

Introducing (2), (3) and (4) in (1), yields

$$dI_p = A(x^2-a^2)^{\frac{1}{2}}\epsilon^{-x^2}x\,dx,$$

where $a^2=(re/2m)^2m/2kT(H^2-H_c^2)$.

We are concerned only with $H>H_c$, i.e., with positive values of $a^2$. Hence $I_p$ will be real only for $x^2>a^2$. Hence we integrate from $a$ to $\infty$,

$$\text{i.e.,}\quad I_p = A\int_a^\infty (x^2-a^2)^{\frac{1}{2}}\epsilon^{-x^2}x\,dx,$$

$$\text{or}\quad I_p = I_1\epsilon^{-a^2} = I_1\epsilon^{(m/2kT)(re/2m)^2(H_c^2-H^2)},$$

$$\text{or}\quad \log\frac{I_p}{I_1} = \frac{m}{2kT}\left(\frac{re}{2m}\right)^2(H_c^2-H^2). \tag{5}$$

$I_1$ in this expression is considered constant, its variation being small compared to that of $I_p$. $I_1$ is the total current emitted by the cathode—not the real cathode but the high temperature virtual cathode discussed previously by the writer.[5] At the cut-off point, i.e., $H=H_c$, it is equal to the space charge limited current $I_0$, and may be computed from the Langmuir-Child law, proper correction being made for electron trajectory curvature due to the magnetic field. For $H>H_c$, the returning current due to those electrons which do not reach the plate and return to the cathode must be considered. This returning current $I_2$ plus the advancing current $I_1$ must approximately equal the space charge limited current, i.e., $I_1+I_2=I_0$. For complete

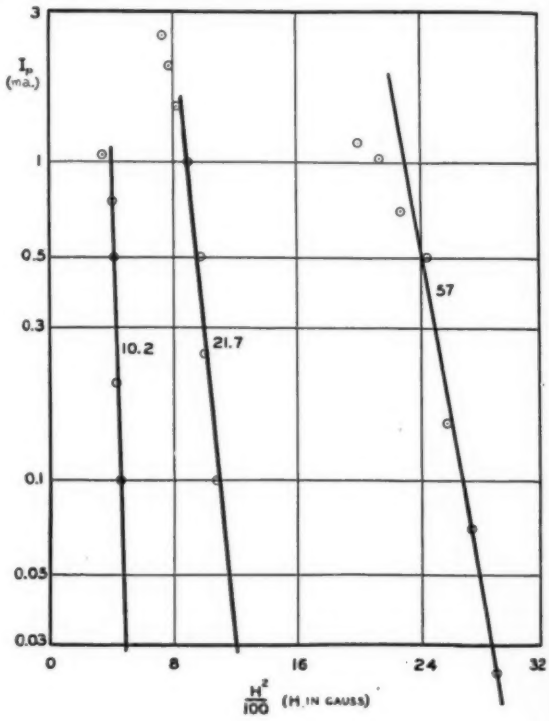



FIG. 2. Log $I_p$ *versus* $H^2$ from Hull's data (reference 1) Fig. 8, page 44.

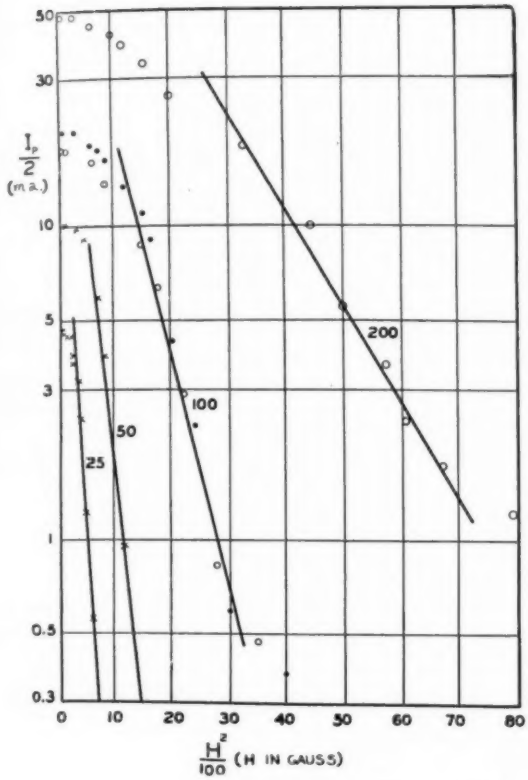



FIG. 3. Log $I_p$ *versus* $H^2$ showing linear portion of curve and change of slope with plate voltage. Curves are labeled with plate voltage $V_p$.

cut-off the plate current becomes zero and we have the returning current equal to the advancing current, i.e., $I_1 = I_2 = I_0/2$. It is thus seen that $I_0/2 \leqslant I_1 \leqslant I_0$, and there is a maximum variation of $I_1$, with respect to $I_0$ of a factor of two.

The space charge limited current $I_0$ itself also may vary as $H$ is increased, due to changes in the size of the electron swarm, and potential distribution. The magnitude of this change may be estimated by taking the Langmuir-Child equation

$$I_0 = \frac{2}{9}\left(\frac{2e}{m}\right)^{\frac{1}{2}} \frac{V_1^{\frac{3}{2}}}{r_1},$$

where $V_1$ is the potential at the edge of the swarm and $r_1$ the radius of the swarm. The magnetic field may be expressed in terms of $V_1$ and $r_1$ by an expression similar to (4)

i.e., $$H = (8mV_1/er_1^2)^{\frac{1}{2}}.$$

Eliminating $r_1$ between these two equations yields $I_0 = (e/9m)V_1H$, which gives the variation of the space charge limited current with $H$ and $V_1$. As $H$ is increased the swarm shrinks in size and therefore $V_1$ decreases. Thus the changes in $H$ and $V_1$ are compensatory. But even without the compensating change in $V_1$ the change in $I_0$ would be negligible compared to that of $I_p$.

The total change in $I_1$ is thus seen to be small compared to the change in $I_p$, and $I_1$ will therefore be assumed constant in discussing the experimental results.

From (5) it is evident, that if log $I_p$ be plotted against $H^2$ a straight line results whose slope depends only upon $T$. Data from Hull's paper plotted in such a fashion are shown in Fig. 2. Similar data obtained by the writer are plotted in Fig. 3. The portions of the curves to which the above analysis applies are fairly linear.

It is evident that the slope, i.e., electron temperature, increases with the plate potential $V_p$. Data from Figs. 2 and 3 showing this are plotted in Fig. 4. The points are somewhat scattered,

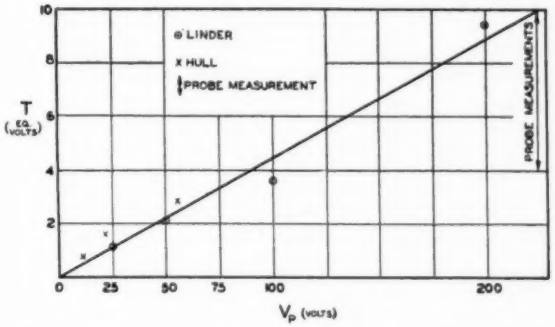



FIG. 4. Electron swarm temperature in equivalent electron volts *versus* plate voltage, as determined from cut-off curve and by probe measurements.

but within the experimental error lie on a straight line through the origin. This is in agreement with previous measurements of electron temperatures by probe methods,[4, 5] in which the electron temperature was found to be proportional to $V_p$ under similar conditions.

Probe measurements were made in the present case also. These revealed that the electron temperature did not remain constant as $H$ varied, but increased as cut-off became more complete. The range of variation is shown by the vertical arrow in Fig. 4. It may be seen that the mean

value given by the probe measurements is in fair agreement with the values determined from the cut-off curves. That the latter points seem somewhat too high may be due to tube dissymmetries which would cause the cut-off curve to have a greater slope, and cause the temperature determined therefrom to be too high.

It may be concluded that the shape of the cut-off curve can be satisfactorily explained on the basis of the random electron motion. The effect of this motion on other magnetron characteristics such as orbit shape, transit time, transit angle, tube noise, etc., appears not to have been considered in any publications up to the present time, but there can be little doubt that these are all affected in an important manner.

---

## The Numerical Solution of Laplace's Equation

G. H. Shortley and R. Weller
*Mendenhall Laboratory of Physics, Ohio State University, Columbus, Ohio*
(Received February 28, 1938)

This paper considers in detail numerical methods of solving Laplace's equation in an arbitrary two-dimensional region with given boundary values. The methods involve the solution of approximating difference equations by iterative procedures. Modifications of the standard Liebmann procedure are developed which lead to a great increase in the convenience and rapidity of obtaining such a numerical solution. These modifications involve the use of formulas which simultaneously improve a block of points in place of a single point; methods of operating on the differences of trial functions in place of the functions themselves; and also a method of extrapolating to the final solution of the difference equations. The theory underlying these procedures is considered in detail by a new method which involves the expansion of the error and difference functions in terms of eigenfunctions. This permits definite comparison of rates of convergence of various procedures. The techniques of handling practical problems are considered in detail.

### §1. Introduction

THERE are many industrial and engineering problems in which solutions of Laplace's equation are needed. These solutions cannot often be obtained analytically because of the occurrence of irregular boundaries for which systems of separable coordinates are unknown. For example, our interest in this equation was aroused by the need for a convenient method of obtaining from the fringe photographs of photoelasticity the principal stresses in the interior of the specimen tested. The photographs give directly only the *difference* of the two principal stresses. On the boundary of the specimen the normal component of the stress vanishes, so here we have complete information. The *sum* of the principal stresses satisfies Laplace's equation; values of this sum in the interior of the specimen can be obtained from the fringe photographs by solving the two-dimensional Laplace equation, using the known boundary values.[1] A numerical method must be used to obtain this solution since the shape of the specimen is usually much too irregular to admit of analytical treatment. The same necessity for numerical solution of Laplace's equation arises in the calculation of electrostatic fields in regions enclosed by boundaries at known potentials; in the calculation of steady state temperature distributions in heat conduction problems; in the calculation of the shapes of films and membranes; and in certain problems of hydrodynamics and gravitation.

We shall restrict our discussion to the solution of the Laplace equation in a two-dimensional

[1] Details of the application to photoelasticity of the theory outlined in this paper will be found in Weller and Shortley, "Calculation of Internal Stresses in Photoelasticity," to be published soon in the *Journal of Applied Mechanics*. This reference should be consulted for an explicitly solved example.

region with given arbitrary boundary values, with some mention of the case where the normal derivative at the boundary is given instead of the boundary value. The methods we shall give can be readily extended to three dimensions, but the amount of labor involved in solving a three-dimensional problem will be tremendously greater than that required for a similar two-dimensional problem.

Any numerical method of handling a differential equation necessarily involves the replacement of the continuum of points on the boundary and in the interior of the region by a discrete set of points. The best method which has been previously given for handling Laplace's equation is an iterative procedure due to Liebmann.[2] In this, a square net is laid down over the region; approximate values are assigned to the function at the interior points—known values at the boundary points. One then traverses the net repeatedly, replacing the value at each interior point by the sum of the values at the four neighboring points, using the new values immediately in the improvement of the succeeding points. The function will then converge to a solution of the Laplace difference equation

$$[\omega(x, y+h)+\omega(x, y-h)+\omega(x+h, y)+\omega(x-h, y)-4\omega(x, y)]/h^2=0. \quad (1)$$

which will be an approximation to the solution of the Laplace differential equation—the smaller the interval $h$, the better the approximation.

[2] An examination of the literature of mathematical and applied physics indicates that numerical methods of solving Laplace's equation have received little attention. Liebmann's paper appears in Sitzungsber. der math.-phys. Klasse der Bayer. Akad., München, 1918, p. 385. The method is discussed by Phillips and Wiener, J. Math. and Phys. **2**, 105 (1923); Courant, Zeits. f. Angew. Math. u. Mech. **6**, 322 (1926); Frank von Mises, *Differential und Integralgleichungen der Mathematischen Physik* I, p. 735. In these references little attention is given to the application of the method to practical problems: none of the material contained in the present paper is discussed. A different successive approximation method for the solution of the difference equation is given by L. F. Richardson, Trans. Roy. Soc. **A210**, 307 (1910); Math. Gazette **12**, 415 (1925). This method is similar to the Liebmann procedure but does not use the new values, as they are obtained, in the improvement of succeeding points. This makes the convergence much slower (cf. section e., Appendix) unless considerable care is exercised in the assignment of certain arbitrary multiplying factors. The method does not have the general utility of the Liebmann, and at its best does not converge as rapidly as the modified methods we shall propose.

It has been shown quite generally by Richardson[2] that the difference between the solution of a differential equation and that of an approximating difference equation on a net of interval $h$ is of the form

$$A(x, y)h^2+B(x, y)h^4+C(x, y)h^6+\cdots. \quad (2)$$

Since only even powers of $h$ enter here, the difference solution approaches the differential solution rapidly as $h$ approaches zero. If $h$ is sufficiently small the error is proportional to $h^2$, and by making two solutions with different $h$, one can estimate the difference of each from the solution of the continuous problem.

In this paper we shall first investigate, by a new method, the rate of convergence of the iterative process to the final solution of the difference equation in the case of the Liebmann procedure. These considerations lead to the formulation of procedures of similar type (in particular the "nine-block" procedure) which converge much more rapidly than the Liebmann. They also lead to the development of an accurate method for extrapolating to the final solution of the difference equations. In this way we greatly reduce the amount of labor required in obtaining a numerical solution of Laplace's equation to any desired accuracy and render the task of calculating the solution entirely practical and convenient. The last section of this paper is devoted to a detailed discussion of practical techniques.

## §2. Theory

In this section we shall for simplicity assume that we are working with a square net of interval $h$ in a region with a rectangular boundary. The slight modifications of the theory needed in case of a more general net or an irregular boundary will be discussed in §3 under the heading of Technique. We shall first examine the rate of convergence of the Liebmann iteration procedure.

Denote the values of a function at the interior points of the lattice by $\psi_i$, where $i=1, 2, \cdots, n$. Denote the values at the boundary points by $\psi_\alpha$, where $\alpha=n+1, n+2, \cdots, m$. For example, for the lattice of Fig. 1, $n=54$, $m=88$.

In the Liebmann procedure for solving the Laplace difference equation (1) with assigned boundary values $\psi_\alpha$ one starts by assuming any

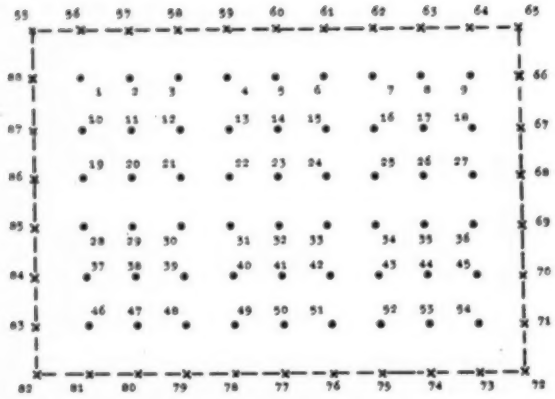


FIG. 1. A net with 54 interior points and 34 boundary points.

values $\psi_i$ whatsoever at the interior points. Then one works over the lattice, e.g. in the order in which the points are numbered in Fig. 1, replacing each $\psi_i$ by the mean of the four neighboring values to get an improved value $\bar{\psi}_i$. *The previously improved values are used in calculating this mean wherever they are available*, so that e.g. in Fig. 1,

$$\bar{\psi}_{13}=\tfrac{1}{4}(\bar{\psi}_{12}+\bar{\psi}_{4}+\psi_{14}+\psi_{22}). \tag{3}$$

The function $\bar{\psi}$ is then closer to the solution of the difference equation than was $\psi$.[3] Iteration of this process leads to values which are as close to the true solution as may be desired.

In this improvement,[4] $\bar{\psi}_i$ may be expressed as a linear combination of all the $\psi_i$ and the $\psi_\alpha$, since for example in (3) the $\bar{\psi}$'s on the right are themselves combinations of other $\bar{\psi}$'s and $\psi$'s which in the last analysis reduce to combinations of $\psi$'s alone. Hence we may write

$$\bar{\psi}_i=\sum_j T_{ij}\psi_j+\sum_\alpha U_{i\alpha}\psi_\alpha, \qquad \bar{\psi}_\alpha=\psi_\alpha, \tag{4}$$

where the $T_{ij}$ and $U_{i\alpha}$ are numerical coefficients which depend on the size of the net and the order of improving the points, but not on the particular boundary values given nor the particular body values assumed.

Now let $\omega$ represent the true solution of the difference equation (1). The *trial function* $\psi$ may then be expressed as the sum of the *true solution* $\omega$ and an *error* $\phi$:

$$\psi_i=\omega_i+\phi_i, \qquad \psi_\alpha=\omega_\alpha+\phi_\alpha=\omega_\alpha. \tag{5}$$

Since the trial function must be chosen with the correct boundary values $\omega_\alpha$, the error $\phi$ will have vanishing boundary values: $\phi_\alpha=0$. Since the transformation (4) is *linear*, the function obtained when $\psi$ is "improved" will be the sum of the functions obtained when $\omega$ and $\phi$ are improved: $\bar{\psi}=\bar{\omega}+\bar{\phi}$. But $\bar{\omega}=\omega$—*the improvement process is such that it does not change the true solution*—this is seen to be true when the improvement formula (3) operates on any solution of (1). Hence

$$\bar{\psi}=\omega+\bar{\phi}, \tag{6}$$

*and in order to study the rate of convergence of an arbitrary function $\psi$ with correct boundary values to the true solution $\omega$ on iteration of the improvement, we need only investigate the rate of convergence of an arbitrary "error" function $\phi$ with zero boundary values to zero as the same improvement formula is used repeatedly.* The two rates of convergence are the same. Since $\phi_\alpha=0$, we have simply

$$\bar{\phi}_i=\sum_j T_{ij}\phi_j, \tag{7}$$

or briefly

$$\bar{\phi}=\boldsymbol{T}\phi. \tag{7'}$$

Now the possible functions $\phi$ are best classified in terms of the eigenfunctions (characteristic vectors) of the matrix $\boldsymbol{T}$. There will be a set of functions $\phi^{(1)}, \phi^{(2)}, \cdots, \phi^{(n)}$ such that

$$\boldsymbol{T}\phi^{(i)}=\lambda_i\phi^{(i)},$$

where $\lambda_i$ is a constant. We show in the Appendix that *all of the $|\lambda|$'s will be less than unity*. This set of eigenfunctions may be considered as forming a complete set[5] in terms of which any function

[3] See the convergence proof in the Appendix, or in the references noted in footnote 2.

[4] The arguments of the next four paragraphs apply equally well to the other types of improvement procedures considered in the balance of the paper and in the Appendix.

[5] The arguments of the text at this point are not rigorously true and therefore the conclusions drawn from them need justification which is furnished in this footnote. The discussion in the text would be mathematically correct only if the matrix $\boldsymbol{T}$ were symmetric, which a calculation for a simple lattice shows to be not the case. For example for a square lattice of four interior points, improved in the order

1 2
3 4

$\phi$ (with zero boundary values) may be expanded:

$$\phi = a_1\phi^{(1)} + a_2\phi^{(2)} + \cdots + a_n\phi^{(n)}. \qquad (13)$$

Then $T\phi = \lambda_1 a_1\phi^{(1)} + \lambda_2 a_2\phi^{(2)} + \cdots$;

$$TT\phi = T^2\phi = \lambda_1^2 a_1\phi^{(1)} + \lambda_2^2 a_2\phi^{(2)} + \cdots,$$

. . . . . .

$$T^k\phi = \lambda_1^k a_1\phi^{(1)} + \lambda_2^k a_2\phi^{(2)} + \cdots.$$

So after $k$ applications of $T$, i.e., after $k$ traverses, the coefficients of the various eigenfunctions in $\phi$ are decreased by $\lambda^k$. After a time only the function with the highest eigenvalue, say $\lambda_1$, will remain. The whole rate of convergence of $\phi$ to zero, and hence of $\psi$ to $\omega$ will depend on the rate at which the $\phi^{(1)}$ part of $\phi$ is wiped out—i.e., on the size of $\lambda_1$.

These considerations furnish a practical method of finding $\lambda_1$ and $\phi^{(1)}$ for a given lattice and order of improvement. $\phi^{(1)}$ will clearly be "smooth," nodeless, and will have a maximum near the center of the region in question. Take a trial function of this general shape, iterate repeatedly, and soon all the terms involving $\phi^{(i)}$'s of lesser eigenvalue will have dropped out, leaving a function proportional to $\phi^{(1)}$, which on each traverse is reduced uniformly to $\lambda_1$ times itself.[6] In this way, the values of $\lambda_1$ for regions with square boundaries of various sizes were determined for the Liebmann procedure. The values are plotted in Fig. 2 in the form of a curve giving $-p^2 \log_e \lambda_1$ as a function of $p$, the number of intervals on a side of the square $[(p-1)^2 = n$,

---

the matrix $T$ is the following nonsymmetric array:

$$\begin{matrix} 0 & 8 & 8 & 0 \\ 0 & 2 & 2 & 8 \\ 0 & 2 & 2 & 8 \\ 0 & 1 & 1 & 4 \end{matrix} \qquad (8)$$

(Here the elements of the top row are $T_{11}$, $T_{12}$, $T_{13}$, $T_{14}$; etc.) A nonsymmetric matrix cannot in general be reduced to diagonal form; its roots will not necessarily be real, and it will not have a sufficient number of *eigenvectors* to permit expansion in terms of them. But it can be reduced to a unique stepmatrix with steps of the form

$$\begin{matrix} \lambda_q & 1 & 0 & \cdot & 0 & 0 \\ 0 & \lambda_q & 1 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \lambda_q & 1 \\ 0 & 0 & 0 & \cdot & 0 & \lambda_q \end{matrix} \qquad (9)$$

Associated with the root $\lambda_q$ there will hence be a group of *invariant vectors* $\chi^{(q1)}$, $\chi^{(q2)}$, $\cdots$ satisfying the relations

$$\begin{aligned} T\chi^{(q1)} &= \lambda_q\chi^{(q1)} \\ T\chi^{(q2)} &= \lambda_q\chi^{(q2)} + \chi^{(q1)} \\ T\chi^{(q3)} &= \lambda_q\chi^{(q3)} + \chi^{(q2)} \\ &\cdots\cdots\cdots; \end{aligned} \qquad (10)$$

of these only $\chi^{(q1)}$ is an eigenvector. The entire set of invariant vectors is complete and the expansion (13) must be made in terms of them. The canonical form of the matrix (8) is

$$\begin{matrix} \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{matrix} \qquad (11)$$

while the canonical form of the matrix for a six-point net improved in the order

$$\begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{matrix}$$

is

$$\begin{matrix} .3643 & 0 & 0 & 0 & 0 & 0 \\ 0 & .0625 & 0 & 0 & 0 & 0 \\ 0 & 0 & .0107 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix} \qquad (12)$$

In spite, however, of the fact that $T$ is not a symmetric matrix, it is convenient to speak in the text as if it were, because of the simplicity and familiarity of the simple eigenvalue theory of symmetric matrices. This is permissible because we are not led into any essential error thereby. It may be shown that repeated application of a stepmatrix composed of steps such as (9) to an arbitrary vector reduces finally to the *eigenvector* of the root of highest absolute value. Thus the conclusions of the text are correct. The rapidity of this convergence to the eigenvector of the highest root is however less, the greater the size of the step involving this particular root. But it seems likely from geometrical and asymptotic considerations that the largest roots are all real and belong to steps of order 1, and that complex roots and invariant vectors which are not eigenvectors occur only for the very small roots, where they are quickly wiped out by the iteration and do not cause trouble. These statements are borne out by the form of (11) and (12), in which the only invariant vectors which are not eigenvectors belong to the root 0. Hence all the text arguments may be said to be "practically" true.

For reference to the theory of nonsymmetric matrices see Bôcher, *Higher Algebra*, Chap. XXI, or Wedderburn, *Matrices*, Chap. III.

[6] The determination of this $\lambda_1$ value by iteration is facilitated by the following theorems, which apply to the improvement of functions of zero boundary value.

1. *If the nowhere negative function $\phi$ goes on improvement into a function $\bar{\phi}$ such that for all lattice points $\mu \leq \bar{\phi}_i/\phi_i \leq \nu$, then $\mu \leq \lambda_1 \leq \nu$.* This theorem follows if we show that the bounds on $\phi/\phi$ imply the same bounds on $T\phi/\phi$, and this follows by noting that in (7) all the $T_{ij}$'s are nonnegative (see Appendix for discussion of this property of $T_{ij}$ in connection with the various types of improvement formulae).

2. When one has arrived at the stage where all the eigenfunctions have disappeared except the first and second, and if $\phi$ is everywhere positive, one will find that along the nodes of $\phi^{(2)}$, $\bar{\phi}_i/\phi_i = \lambda_1$ and is constant on iteration; in regions where $\phi^{(2)} > 0$, $\bar{\phi}_i/\phi_i < \lambda_1$ but increases and approaches $\lambda_1$ with iteration; in regions where $\phi^{(2)} < 0$, $\bar{\phi}_i/\phi_i > \lambda_1$ but decreases and approaches $\lambda_1$. To prove these statements consider a function of the form $\phi = \phi^{(1)} + \phi^{(2)}$. On iteration $\phi \rightarrow \bar{\phi} = T\phi \rightarrow T\bar{\phi}$. Investigation of the successive ratios, $T\phi/\phi$ and $T\bar{\phi}/\bar{\phi}$ gives the above results when we remember that $\lambda_2 < \lambda_1$. Hence *when one arrives at the point where the ratio $\bar{\phi}/\phi$ between successive functions is decreasing with iteration over roughly half the region, increasing over the other half, and standing constant along the line dividing these regions, the value of the ratio along this dividing line* is $\lambda_1$. By making use of this theorem one need not wait for $\phi^{(2)}$ to disappear but only for $\phi^{(3)}$, $\cdots$, to disappear in determining $\lambda_1$.

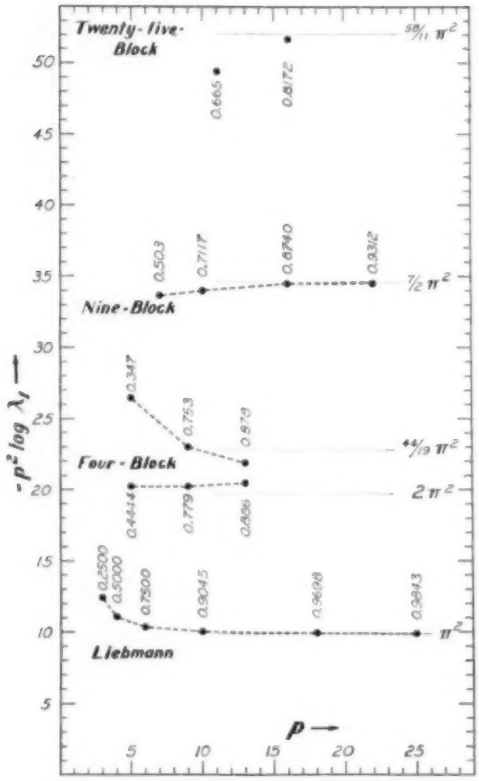



FIG. 2. Relative rates of convergence of the different procedures. The numbers opposite the points are the values of $\lambda_1$ for a square region with $p$ intervals on a side.

the total number of interior points]. This form was chosen because asymptotically for large $p$, $-p^2 \log \lambda_1 \rightarrow \pi^2$, i.e. for large $p$

$$-\log \lambda_1 = 1-\lambda_1 = \pi^2/p^2. \qquad (14a)^7$$

[7] In the case of a rectangular region of $p$ by $q$ intervals, this asymptotic formula becomes

$$-\log \lambda_1 = 1-\lambda_1 = \tfrac{1}{2}\pi^2\left(\frac{1}{p^2}+\frac{1}{q^2}\right). \qquad (14b)$$

A lower eigenvalue $\lambda_{\mu,\nu}$, which belongs to the function with $\mu-1$ nodes in one direction and $\nu-1$ nodes in the other, has the same asymptotic value as that for the nodeless eigenfunction in a region of $p/\mu$ by $q/\nu$ intervals. Hence asymptotically

$$-\log \lambda_{\mu,\nu} = 1-\lambda_{\mu,\nu} = \tfrac{1}{2}\pi^2\left(\frac{\mu^2}{p^2}+\frac{\nu^2}{q^2}\right). \qquad (14c)$$

The derivation of these asymptotic relations is discussed in the Appendix. Although for small regions the form of $\phi^{(1)}$ changes considerably with a change in the order of improvement of the points, the value of $\lambda_1$ is to a good approximation independent of this order. The curve of Fig. 2 was obtained by the use of similar orders for the various size regions, but for all practical purposes it is valid for any order. No particular order seems to have appreciable advantage over any other. While we have no rigorous proof, we believe the asymptotic value to be independent of the order of improvement.

The number of traverses which must be made in order to reduce the coefficient of $\phi^{(1)}$ in (13) to a definite fraction $\rho$ of its initial value is $(-\log \rho)/(-\log \lambda_1)$, so the *value of* $-\log \lambda_1$ *is a measure of the rapidity of convergence of the process.* Approximately, the rapidity of convergence varies inversely as the number of points in the square.

We can then visualize the improvement process as follows: after a number of traverses have been made, the error will be a pillow-shaped multiple of $\phi^{(1)}$. A traverse consists in going over this pillow bringing successive points down to the mean level of their four neighbors. In this way the height of the whole pillow is gradually reduced to zero. We then see that if we could reduce at one step a greater area of the pillow to the "mean" level of its neighbors, the pillow would disappear faster—we would be cutting it down by larger slices. This more rapid slicing would be accomplished if we divided the interior of the region up into square blocks each containing $v^2$ points and then used improvement formulas which would make a whole block at a time a harmonic function [in the sense of (1)] with boundary values determined by the values of the function at the $4v$ points just outside the block. Such improvement formulas can be obtained by solving the $v^2$ simultaneous equations which state that the value at each point in the block equals the mean of its four neighbors.

For example, we might divide the region of Fig. 1 up into blocks of nine points (nine-blocks) in the manner indicated in the drawing. We could then improve the six blocks successively in any order, for example the top row from left to right, then the bottom row from left to right. In this case, the improved values in the upper left-hand block would be calculated from the boundary values $\psi_{56}$, $\psi_{57}$, $\psi_{58}$, $\psi_4$, $\psi_{13}$, $\psi_{22}$, $\psi_{30}$, $\psi_{29}$, $\psi_{28}$, $\psi_{86}$, $\psi_{87}$, $\psi_{88}$. The simplest formulas which give harmonic improved values are

$$\bar{\psi}_2 = \tfrac{1}{112}(37\psi_{57}+11\psi_{56}+11\psi_{58}+11\psi_{88}+11\psi_4 +7\psi_{87}+7\psi_{13}+3\psi_{86}+3\psi_{22}+3\psi_{28}+3\psi_{30}+5\psi_{29});$$

$\bar{\psi}_{10}$, $\bar{\psi}_{12}$, $\bar{\psi}_{20}$ are given by corresponding formulas;

$$\bar{\psi}_1 = \tfrac{1}{4}(\psi_{56}+\psi_{88}+\bar{\psi}_2+\bar{\psi}_{10}); \qquad (15)$$

$\bar{\psi}_3$, $\bar{\psi}_{19}$, $\bar{\psi}_{21}$ are given by corresponding formulas;

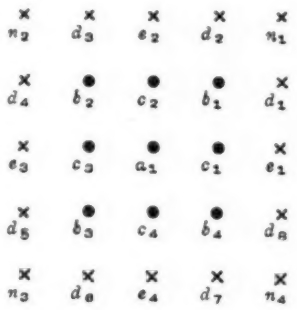
$$\bar{\psi}_{11} = \tfrac{1}{4}(\bar{\psi}_2+\bar{\psi}_{10}+\bar{\psi}_{12}+\bar{\psi}_{20}).$$

In the improvement of the next block, $\bar{\psi}_3$, $\bar{\psi}_{12}$, $\bar{\psi}_{21}$ would be used as three of the boundary values.

By improving by blocks in this way one obtains considerably lower $\lambda$ values and hence convergence to the final solution after fewer traverses than for the straight Liebmann process. We show in the Appendix that the asymptotic values of $-p^2 \log \lambda$ for a $v^2$-block are $v$ times the values given by (14), so the convergence is $v$ times as fast as for the Liebmann. But the formulas used are so complex that it takes $v$ times as much labor per traverse, so that little saving in work is made in this way.

### Nine-block procedure

But it is possible to modify the improvement formulas[8] so that the improvement can be by nine-blocks with practically no more work per improvement than in the case of the Liebmann. In the nine-block




we require that the value

$$a_1=\tfrac{1}{16}[d_1+d_2+d_3+d_4+d_5+d_6+d_7+d_8) \\ +2(e_1+e_2+e_3+e_4)]$$

as would be obtained by using (15); but we may now obtain the values $b_i$ by requiring that they be the means of their four diagonal neighbors, i.e. that

$$b_1=\tfrac{1}{4}(a_1+e_1+e_2+n_1), \quad \text{etc.},^{9}$$

and finally that the $c$ points be the means of their immediate neighbors:

$$c_1=\tfrac{1}{4}(a_1+b_1+e_1+b_4), \quad \text{etc.},$$

using in the last two formulas the improved $a$ and $b$ values.

Precisely the same reasoning which we have applied to the Liebmann procedure may be used in the study of the rate of convergence of this nine-block procedure; empirical values of $\phi^{(1)}$ and $\lambda_1$ for nets of various sizes may be similarly obtained, except that in this case the disappearance of all functions below $\phi^{(1)}$ will take place much more rapidly since all eigenvalues will be much smaller. The values of $-p^2 \log \lambda_1$ obtained in this way are plotted in Fig. 2—they are seen to be about 3.5 times the corresponding values for the Liebmann procedure. The asymptotic value of $-p^2 \log \lambda_1$ for this procedure is $7\frac{1}{2}\pi^2=34.5$ and other asymptotic eigenvalues for this procedure are given by writing $7\frac{1}{2}\pi^2$ in place of $\pi^2$ in (14) (see Appendix). These values are essentially independent of the order of improving the blocks.

Since the number of traverses necessary to accomplish a certain reduction in the coefficient of an eigenfunction in the error is inversely proportional to $-\log \lambda_1$, the number of times the Liebmann procedure must be iterated in order to accomplish reduction by the same factor as in one traverse by the nine-block procedure is just the ratio of the curves of Fig. 2. Thus we may conclude that in general *the nine-block procedure will cut the time required to solve a problem to given accuracy by a factor of almost* 3.5 *as compared to the straight Liebmann method*. This remains true when an extrapolation procedure to be discussed later is used to remove the first eigenfunction after the second has disappeared. The slightly greater inconvenience in improvement of the center-point of the nine-block is made up by a certain advantage in this extrapolation procedure.

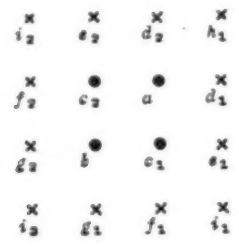
[8] This modification amounts to a change in the set of difference equations which are satisfied by the final solution, so that the final solution for this procedure will be slightly different from that for the Liebmann. That this does not in general impair the accuracy of approximation to the continuous solution is indicated by arguments in Section c of the Appendix.

[9] In case one of the $n$ points is not available, the following formula, derived by the methods outlined in the Appendix, may be used for the corresponding $b$ point:

$$b_1=\tfrac{1}{8}(2a_1+e_1+2d_1+2d_2+e_2).$$

### Four-block procedure

It is possible to use blocks of four in place of blocks of nine, by the following formulas:




$$a = \tfrac{1}{24}(7d_1+7d_2+2e_1+2e_2+2f_1+2f_2+g_1+g_2),$$
$$b = \tfrac{1}{4}(a+f_1+f_2+i_2),$$
$$c_1 = \tfrac{1}{4}(a+b+e_1+f_1), \text{ etc.}$$

The formulas should be used in this order, using the improved values of $a$ and $b$ as they are obtained. The value of $\lambda_1$ varies according to which of the four points is taken as the point $a$ to which to apply the long formula, in relation to the order of improvement. The upper curve of Fig. 2 is for the case in which the long formula is applied to the point of each four block which is nearest the boundary of the region, the lower curve when this formula is applied to the point nearest the center—improving the blocks row by row as one reads the words of a book. The two asymptotic values given are for this same order of improvement, the higher one, $44/19\pi^2$, when the long formula is applied to one of the points on the left side of each block, the lower, $2\pi^2$, when this formula is applied to a point on the right side.

One sees that the four-block converges somewhat more than twice as fast as the straight Liebmann, and is a definite saving in time, in spite of the fact that the labor per improvement is even more than that for the nine-block.

### Twenty-five-block procedure

In the direction of blocks of size larger than nine, one could use a block of 16; in this case it would be necessary to find all four central values directly from the boundary values by means of fairly cumbersome formulae. One can do a twenty-five-block almost as easily, calculating only four points by cumbersome formulae (see Fig. 3). One first obtains the values at the points $a_1$, $a_2$, $a_3$, $a_4$ by formulas similar to the following:

$$a_1 = \tfrac{1}{528}(38d_1+38d_2+76e_1+76e_2+49f_1+49f_2+26g_1+26g_2 +11h_1+11h_2+11i_1+11i_2+18j_1+18j_2+17k_1+17k_2 +12l_1+12l_2+6m_1+6m_2).$$

Then the values at the points marked $b$ are obtained from these and the boundary values by taking the mean of the four diagonal neighbors,[9] and finally the values at the points marked $c$ are obtained from the means of the four closest neighbors. This procedure gives an asymptotic value $-p^2 \log \lambda_1 = 58/11\pi^2 = 52.0$, and hence considerably more rapid convergence than the nine-block. But it is probably quicker on the whole to introduce only nine-blocks since they are so much simpler and easier to improve.

One compensating factor in favor of the 25-block is that the central $b$ and the four central $c$ values need not be calculated until the iteration is completely finished since they do not enter into the improvement formula for any other points.

It is of course possible to handle still larger blocks by similar methods with even lower $\lambda_1$ values, but in these cases so many points would need very complicated improvement formulas that they would certainly not save time on the whole. The optimum block size seems to be nine, with perhaps some arguments in favor of 25.

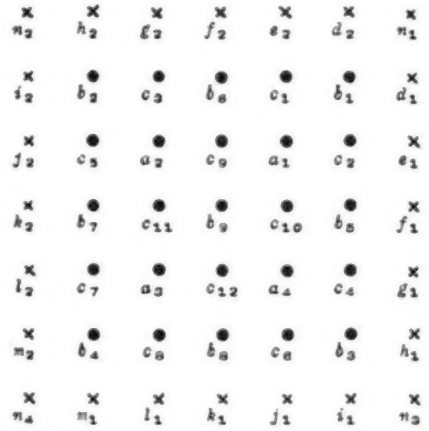



FIG. 3.

## §3. TECHNIQUE

In this section we shall summarize the practical techniques which our experience has shown to be best adapted to the solution of the problem in hand.

### Choice of net

The first thing of course is to lay over the region in question a lattice. It is usually most convenient to take a square lattice, although there are probably cases where a rectangular lattice is advisable, and cases where it is advisable to change the lattice spacing near sections of the boundary where the function representing the boundary values has exceptionally high or low curvature.[10] The net points at which the value of the function is to be defined are the interior points at the intersections of the bars of the lattice and the boundary points at the intersection of the lattice bars with the boundary of the region, as indicated in Fig. 4.

[10] See Appendix for improvement formulas for these cases.

*It is advisable always to start by solving the difference equations with a very coarse net and then successively cut the net spacing in half.* The previous solutions of the coarser nets are desirable because at each halving of the net constant, the convergence is about four times as slow (cf. Fig. 2). Hence it is economical of time to get as good a function as possible from a coarse net to use as the initial trial function for the finer net, so as to keep the error in this function as small as possible.

*The net should be split up into as many nine-blocks as possible and the remainder into as many four-blocks as possible.* Each nine-block or four-block must be surrounded by points at the regular net spacing. These 12 or 8 surrounding points will ordinarily be interior points of the net but may, if a lattice line coincides with a section of the boundary, be boundary points, as at the extreme right of Fig. 4. It seems advisable to observe these requirements rather than to go to the trouble of using formulas modified to take care of unusual distances of the points surrounding the blocks.[9] The more nine- and four-blocks one can use, the lower will be the eigenvalues and the faster the convergence. It occurs to one that it might be advantageous to use blocks which overlap, but we find that the extra labor involved with such blocks more than offsets the gain in rate of convergence.

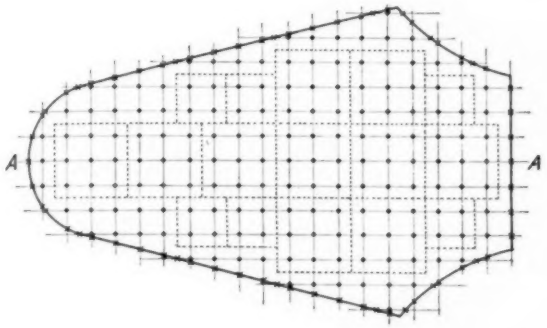



FIG. 4. Location of interior points (circles) and boundary points (crosses).

The reduction of the net spacing should be stopped when one has achieved the desired accuracy in the solution, which may of course not be greater than the accuracy with which the boundary values are known. The amount by which the solution of the difference equations differs from that of the differential equation may be judged from the changes in final solutions for successively finer nets [cf. (2), which however may be expected to apply to the block procedures only on the average since the mode of division of successive nets into blocks will be different].[11]

## Choice of trial function

*The values of the function at the boundary points* are of course determined by the given function which determines the boundary values for the region. In case this function is given only in the shape of a table of values at a discrete set of points around the boundary, the values at the boundary net-points must be determined by interpolation in this table, which is usually done most conveniently by plotting the given boundary values in some convenient way and reading off the values at the net points.

The coarsest net used will usually converge so rapidly that one need not exercise particular care in guessing trial values for the interior points. When the net spacing is halved, the new trial function has the values of the final solution of the coarse net at the points which the two nets have in common and values interpolated from this solution at the remainder of the points. *This interpolation may be very conveniently done by formulas which guarantee that Laplace's equation will be approximately satisfied at these points:* If in Fig. 5 the black dots represent the points of the coarse net, the letters points at which interpolated values are necessary, one takes the value at a point marked $a$ as the mean of the values at the four dots diagonally surrounding $a$, and then the value at a point marked $b$ as the

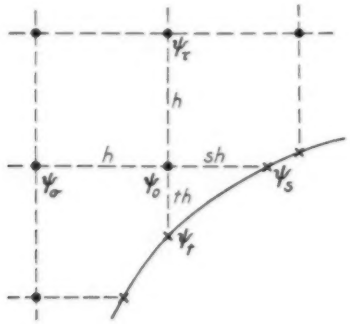
[11] Some indication of the error in the difference solution may be obtained by comparing, at various points, the average of the four nearest neighbors and the average of the four diagonal neighbors. The error is certainly greater than the difference between these two averages. For the Liebmann procedure in a square net of $p \times p$ intervals, we can show that asymptotically the error is at most of the order of $0.06\ p^2$ times the difference between the value at a point and the average of its four diagonal neighbors. This is the case where the difference between the value at a point and the average of its four neighbors for the continuous solution is constant over the net. If the continuous solution is less regular the factor 0.06 is to be decreased. Such definite quantitative statements are not easy to make for other procedures.

mean of the values at the four nearest neighbors, two dots and two $a$ points.

FIG. 5. Interpolation.

### Improvement of points near boundary

An interior point which has boundary values for neighbors will ordinarily be surrounded by two interior points at the regular net spacing $h$




and two boundary points which are closer than $h$, say at distances $sh$ and $th$. However it may be surrounded by three interior points, in which case either $s$ or $t$ may be taken as unity. The improved value of $\psi_0$ in this case is chosen so that Laplace's equation is approximately satisfied at the five points in question. The formula for this improved value, which is a special case of Eq. (23) of the Appendix, is

$$\psi_0=\frac{st}{(s+t)}\left\{\frac{1}{s(1+s)}\psi_s+\frac{1}{1+s}\psi_\sigma+\frac{1}{t(1+t)}\psi_t+\frac{1}{1+t}\psi_\tau\right\}. \quad (16)$$

The complete value of the coefficients of $\psi_\sigma$ and $\psi_\tau$ and the whole terms in $\psi_s$ and $\psi_t$ if these are constant boundary values should be calculated once for all and written on the net. The value $\psi_0$ is then obtained very simply by multiplying $\psi_\sigma$ and $\psi_\tau$ by known coefficients and adding a known constant.

### Axes of symmetry

If there is an axis of symmetry such that the function values are known to be the same at corresponding points on opposite sides of this axis ($AA$ might be such an axis in Fig. 4), it is only necessary to work with the net on one side of this axis. One imagines that each time a value is changed, the value at the corresponding point is simultaneously changed, and in improving values near or on the symmetry axis, one uses these imaginary values as needed. For example, if the left-hand row of points in Fig. 6 (p. 346) lie along a symmetry axis, and one is using the simple Liebmann formula for these points, the improvement formula for $r$ is

$$\psi_r=\tfrac{1}{4}(\psi_s+2\psi_t+\psi_u).$$

One obtains somewhat more rapid convergence in the Liebmann case if the net is chosen so that the symmetry axis is along one of the rows of net points rather than between two rows, and in the block procedures, if the blocks straddle the symmetry axis rather than be located unsymmetrically.

In a similar fashion one can handle a straight section of the boundary where the boundary condition is that the normal derivative of $\psi$ have the value zero, or even have given nonzero values.

### Use of the difference function

In most cases the labor of improvement will be materially reduced if one obtains the difference between the first improved value and the original value, and then applies the improvement formula to this difference function. The difference function is easier to handle because it usually has many fewer significant figures than the original function, because it has zero boundary values, and because it permits of convenient extrapolation. In the notation of (5), (6), (7),

$$\psi=\omega+\phi:$$

the trial function equals the true solution plus

the error; when this is improved, one obtains

$$\bar{\psi} = \omega + T\phi.$$

The difference between $\bar{\psi}$ and $\psi$ we call the difference function

$$\delta = \bar{\psi} - \psi = T\phi - \phi.$$

Since this has zero boundary values, if we apply the improvement process to the difference function we obtain successively the functions

$$\begin{aligned} T\delta &= T^2\phi - T\phi, \\ T^2\delta &= T^3\phi - T^2\phi, \\ &\cdots\cdots\cdots \\ T^{n-1}\delta &= T^n\phi - T^{n-1}\phi. \end{aligned} \tag{17}$$

If we add the sum of these functions to $\bar{\psi}$, we obtain

$$\bar{\psi} + T\delta + T^2\delta + \cdots + T^{n-1}\delta = \omega + T^n\phi, \tag{18}$$

which is the same function we would have obtained by improving $\psi$ itself $n$ times.

### Extrapolation

The greatest advantage of using the difference function lies in the fact that it is possible at a certain point to make an accurate extrapolation to find the sum of the infinite series of terms of the type (17) which must be added to $\bar{\psi}$ to obtain the true solution $\omega$. [If $n = \infty$ in (18),

$$\bar{\psi} + \sum_{i=1}^{\infty} T^i\delta = \omega,$$

since $\phi$ disappears after an infinite number of applications of $T$.] Since $\delta$ has vanishing boundary values, it may be expanded in the series, cf. (13),

$$\delta = b_1\phi^{(1)} + b_2\phi^{(2)} + \cdots + b_n\phi^{(n)}. \tag{19}$$

As $T$ is applied successively to $\delta$, the functions $\phi^{(2)}, \cdots, \phi^{(n)}$ disappear before $\phi^{(1)}$, and after a time $T^i\delta$ will become a fairly smooth, pillow-shaped, one-signed function decreasing uniformly by a factor $\lambda_1$ at each iteration. At this stage, one has

$$\begin{aligned} T^k\delta &= c\phi^{(1)}, \\ T^{k+1}\delta &= \lambda_1 c\phi^{(1)}, \\ T^{k+2}\delta &= \lambda_1^2 c\phi^{(1)}, \\ &\cdots\cdots\cdots \end{aligned}$$

so that

$$\begin{aligned} &T^k\delta + T^{k+1}\delta + T^{k+2}\delta + \cdots \\ &\quad = c\phi^{(1)}(1 + \lambda_1 + \lambda_1^2 + \cdots) = T^k\delta/(1 - \lambda_1), \end{aligned}$$

which furnishes the evaluation of the desired sum. The time saved by this extrapolation depends on the relative rate of convergence of the first and second eigenfunctions—after the second eigenfunction has been iterated out of $\delta$, the first may be extrapolated out.[12]

### Resumé

We conclude with a detailed resumé of the calculation for a given net, making use of the difference function and extrapolation.

Having chosen the trial function, the improvement formula to be used for each point, and the order of going over the net, *go over the trial function $\psi$ once to obtain the function $\bar{\psi}$.*

*Calculate the difference function* $\delta = \bar{\psi} - \psi$. This will have zero boundary values. It need not be computed at points whose *unimproved* function values are not used in the improvement of any other point. In the nine-block procedure, this usually applies to four of the nine points in each block, e.g. to points 41, 42, 50, 51 in the bottom center block of Fig. 1 if the blocks are improved in the order of reading a book.

*Improve the function $\delta$ repeatedly* in the same order and with the same improvement formulas as used for $\psi$, to obtain successively $T\delta$, $T^2\delta$, $T^3\delta$, $\cdots$. When this function has become fairly smooth, one-signed,[13] and falling approximately

[12] For a large net in a boundary approximately square or circular, the extrapolation will reduce the number of necessary traverses by a factor of 2.5 since $\log \lambda_2/\log \lambda_1 \approx 2.5$ in this case, cf. (14c). This factor is reduced as the region becomes more rectangular or elongated, dropping to about 1.3 for a rectangle 3 times as long as broad. But the factor is increased by the presence of axes of symmetry since in such cases the second eigenfunction will have a node along the axis of symmetry and the fact that all the successive approximation functions are automatically symmetrical means that the coefficient of this second eigenfunction in an expansion such as (19) is identically zero; in this case one will be ready for extrapolation as soon as the third eigenfunction has disappeared. For a region approximately square or circular with two crossing symmetry axes, extrapolation reduces the labor by a factor of about 5. For a rectangular region with one axis of symmetry across the short way, the factor is a maximum of 3.2, falling to 1.8 for a rectangle 3 times as long as broad.

[13] Cases sometimes occur where $T^k\delta$ is two-signed with an approximately stationary node and looks more like the second eigenfunction than the first. In this case it is profitable to extrapolate at this stage.

at a uniform rate, calculate the ratio $T^k\delta/T^{k-1}\delta$ at various points of the net. If the higher eigenfunctions were completely eliminated, this ratio would be constant and equal to $\lambda_1$. It does not pay to wait for this to become strictly true before extrapolating. Fig. 2 and the discussion of footnote 6 assist in finding an early approximation to the value of $\lambda_1$.

Using an approximate $\lambda_1$ determined in this way, *add together* $\bar{\psi}$, $T\delta$, $T^2\delta$, $\cdots$, $T^{k-1}\delta$, $T^k\delta/(1-\lambda_1)$ *to obtain a new trial function* $\psi'$. *Apply the improvement process to* $\psi'$ *to obtain* $\bar{\psi}'$. ($\psi'$ *need be determined only at those points at which its value is used in determining* $\bar{\psi}'$. In the case of the nine-block procedure this fact makes a considerable saving in the time of extrapolation.) The function $\psi'$ should be very close to the final function $\omega$, its departure from $\omega$ is measured by the difference between $\psi'$ and $\bar{\psi}'$. If necessary, *calculate a new difference function* $\delta'=\bar{\psi}'-\psi'$ *and repeat the procedure*. The more constant one requires the ratio $T^k\delta/T^{k-1}\delta$ to be before extrapolating the better will be the function $\psi'$, but it seems advisable to make a fairly early extrapolation, and then continue the improvement, for two reasons. First, the new $\delta'$ will in general be much smaller than $T^k\delta$, and hence it will be much easier to improve it than to continue improving $T^k\delta$. Second, there is a possibility that one may have made an error in obtaining $\bar{\psi}$ or one of the $T^i\delta$'s, and if this has occurred, it will show up in $\psi'$, so it is advisable to make a fresh start before one does too much work, just as a check.

In order to obtain a certain number of decimal places in the final solution $\omega$, it is necessary to carry more places in $\bar{\psi}$ and $\delta$ in order to avoid large rounding-off errors. It is necessary that a change of $1/(1-\lambda_1)$ in the last figure carried in $\delta$ represent a change of less than 1 in the place where the last significant figure of $\omega$ is to be located. It is desirable for ease of calculation of $\lambda_1$ that the number of figures in $\delta$ should be chosen more generously than this. For a net not exceeding a couple of hundred points, *it is convenient to carry two more figures in* $\bar{\psi}$ *and* $\delta$ *than one wants in the final solution*. One then need carry the process only to the point where the estimated error of extrapolation is less than 50 in the last decimal place to get an error of less than $\frac{1}{2}$ in the last place of $\omega$.

With regard to the time and labor involved in the numerical solution of problems by these methods, no general statement can be made since so much depends upon the accuracy required. To engineering accuracy of one percent a reasonably complicated problem such as arises in photoelasticity can be solved in two or three hours, after the boundary values have been determined.[1] For problems requiring greater accuracy we feel that the method is sufficiently convenient to be entirely practical and deserving of wide application to physical problems.

We believe that methods similar to those discussed above can be applied to the solution of Poisson's equation $\Delta\psi=\rho(x, y)$. We are planning to consider the numerical solution of this equation in the near future.[14]

We are indebted to Professor L. H. Thomas for helpful discussions concerning approximation formulas.

[14] The heat conduction and vibration equation, $\Delta\psi=\lambda\psi$ and Schrödinger's equation $\Delta\psi+\rho(x,y)\psi=\lambda\psi$ are more complicated because an eigenvalue $\lambda$ must be determined at the same time that $\psi$ is determined. A generalization of the Liebmann procedure suitable for this type of equation has been considered by Kimball and Shortley, Phys. Rev. **45**, 815 (1934).

## Appendix

### a. Approximation formulas

We shall derive first the series which gives the value of a harmonic function $\psi(x,y)$ at a given point (which we may conveniently take to be the origin 0,0) in terms of the function values $\psi_1, \psi_2, \cdots, \psi_n$ at selected neighboring points $x_1, y_1; x_2, y_2; \cdots; x_n, y_n$. This will enable us to obtain improvement formulas for points of a net which are not surrounded by four regularly disposed points. In the neighborhood of the origin, any harmonic function can be represented by the series

$$\psi(x,y)=\alpha_0+\alpha_1 r\cos\theta+\alpha_2 r^2\cos 2\theta+\cdots +\beta_1 r\sin\theta+\beta_2 r^2\sin 2\theta+\cdots, \tag{20}$$

where we have introduced polar coordinates $r$ and $\theta$; $x=r\cos\theta$, $y=r\sin\theta$. The terms $r^n\cos n\theta$ and $r^n\sin n\theta$ here are homogeneous $n$th order polynomials in $x$ and $y$, which are the real and imaginary parts of $(x+iy)^n$.

We now wish to approximate the value $\psi(0,0)=\psi_0$ of the arbitrary harmonic function (20) in terms of the function values $\psi_1, \psi_2, \cdots, \psi_n$. We assume it to be the following linear combination

$$\psi_0=\sum_{i=1}^{n} b_i\psi_i. \tag{21}$$

When we substitute from (20) the values of these $\psi$'s, this equation becomes

$$\alpha_0=\alpha_0\Sigma b_i+\alpha_1\Sigma b_i r_i\cos\theta_i+\alpha_2\Sigma b_i r_i^2\cos 2\theta_i+\cdots +\beta_1\Sigma b_i r_i\sin\theta_i+\beta_2\Sigma b_i r_i^2\sin 2\theta_i+\cdots.$$

Since the $\alpha$'s and $\beta$'s are to be arbitrary, we equate their coefficients to zero to obtain the following set of equations to determine the $b$'s of (21):

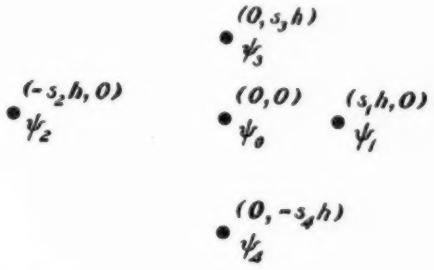
$$\begin{aligned}
&\Sigma b_i=1,\\
&\Sigma b_i r_i\cos\theta_i=\Sigma b_i x_i=0,\\
&\Sigma b_i r_i\sin\theta_i=\Sigma b_i y_i=0,\\
&\Sigma b_i r_i^2\cos 2\theta_i=\Sigma b_i(x_i^2-y_i^2)=0,\\
&\Sigma b_i r_i^2\sin 2\theta_i=\Sigma b_i x_i y_i=0,\\
&\cdots\cdots\cdots\cdots
\end{aligned} \tag{22}$$

Of course, since the number of $b$'s is only $n$, one cannot satisfy this whole infinity of equations. One chooses the $b$'s to satisfy as many consecutive equations, beginning at the top, as possible; this corresponds to putting a harmonic function like (20) of as high an order as possible through the given points.

It is Eq. (21) which is used as one of the set of difference equations which replace the Laplace differential equation, and which is also used for the improvement formula for the point in question.

If $n=5$, we can in general satisfy the first five of equations (22) and hence make use of a quadric interpolation function. The equation of footnote 9 is an example of this case.

If $n=4$, we cannot in general find a set of $b$'s satisfying the first five equations of (22), so that we cannot use a complete quadric interpolation formula. If however the four points are located on two straight lines passing through the origin and intersecting at 90°, there is a set of $b$'s which satisfy the first five equations of (22). We see this if we take these lines as axes of coordinates, in which case the fifth equation of (22) is automatically satisfied, and the same $b$ values clearly satisfy the equations when each $\theta_i$ is decreased by $\alpha$, corresponding to a rotation of coordinate axes through angle $\alpha$. One of the coefficients $\alpha_2$ or $\beta_2$ is undetermined by the function values at the four points, but all harmonic quadrics through the four points have the same value at the origin. For four points located at distances $s_1h$, $s_2h$, $s_3h$, $s_4h$ from the origin, as in the figure,

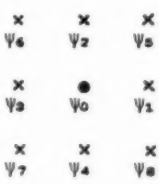



the value of $\psi_0$ is given by

$$\left(\frac{1}{s_1 s_2}+\frac{1}{s_3 s_4}\right)\psi_0=\frac{1}{s_1(s_1+s_2)}\psi_1+\frac{1}{s_2(s_1+s_2)}\psi_2 +\frac{1}{s_3(s_3+s_4)}\psi_3+\frac{1}{s_4(s_3+s_4)}\psi_4. \tag{23}$$

The improvement formula for points near a boundary (16) and points of a *rectangular* net with unequal spacing in the two directions follow immediately from this formula. For the special case $s_1=s_2=s_3=s_4=1$, where $\psi_0$ is just the average of the four neighboring points, the first 7 equations of (22) are satisfied, so this formula is good to the third order—for all cubic harmonic polynomials, the value at a point is the mean of four equidistant neighbors.

For $n>5$ we can in general satisfy Laplace's equation to a higher order. For example if we take 8 points arranged in a square we find that




$$20\psi_0=4(\psi_1+\psi_2+\psi_3+\psi_4)+\psi_5+\psi_6+\psi_7+\psi_8. \tag{24}$$

The $b$ values here satisfy the first *fifteen* equations of (22) so that Laplace's equation is satisfied to the *seventh* order by this formula.

### b. Change of net constant

The considerations of the preceding section enable us to find suitable improvement formulas for a region such as that shown in Fig. 6 where the interval is being changed

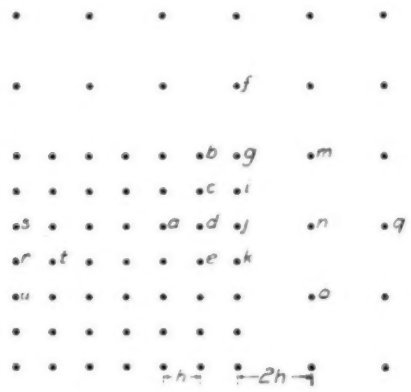



FIG. 6. Change of net spacing.

from the value $h$ in one part of the net to double this value in another. Such a refinement of the net is sometimes desirable near a region of the boundary where the function curves very rapidly. Typical formulas are

$$\psi_d = \tfrac{1}{4}(\psi_a+\psi_c+\psi_j+\psi_e),$$
$$\psi_n = \tfrac{1}{4}(\psi_j+\psi_m+\psi_q+\psi_o),$$
$$\psi_g = \tfrac{1}{6}(2\psi_b+\psi_f+\psi_m+2\psi_i),$$
$$\psi_j = \tfrac{1}{6}(2\psi_d+3\psi_i+\psi_n+3\psi_k),$$
$$\psi_i = \tfrac{1}{16}(4\psi_e+5\psi_g+\psi_m+\psi_n+5\psi_j).$$

### c. Accuracy of the difference solution. The seventh-order procedure

We wish here to make a few general statements concerning the accuracy with which the converged solution obtained for a given net approximates to the continuous solution. In the Liebmann procedure the value of the function at a point is given in terms of the neighboring values correct to third-order terms. But the formula (24) satisfies the equation in the square surrounding each point to the seventh order. Hence in general one might expect that the use of (24) as an improvement formula will give a sufficiently closer approximation to the continuous solution than the other methods we have discussed to justify its use in spite of the extra labor entailed. The labor is greater because about three times the work per point is required in the improvement over that in the Liebmann or nine-block, while the rate of convergence is but slightly faster than that of the Liebmann and only $\frac{1}{3}$ as fast as the nine-block. (The asymptotic value for the seventh-order process, as we shall call this procedure, of $-p^2 \log \lambda_1$ is found to be $\frac{6}{5}\pi^2$ for a square net.) But in most engineering problems it is doubtful that one gains by using so accurate a formula, since it corresponds to using an interpolation formula which is more accurate than is indicated by the smoothness of the boundary values. Also, in the case of an irregular boundary, it is very inconvenient to attempt to use improvement formulas which are more accurate than the second-order accuracy of (16) for points near the boundary.

In cases where great accuracy is desired in the solution and is justified by the gift of a boundary function known analytically or to high accuracy and for which at least the fourth differences are smooth, the seventh-order formula would undoubtedly be very useful. To see how great a gain in accuracy is made by use of this formula in such a case, we chose an example in which the boundary values were taken so that the solution of the differential equation was $\log_{10} r^2$ in the region with boundaries $x=100$ and 110; $y=-5$ and $+5$; net constant $h=1$ (81 interior points). In this case the maximum discrepancy from the continuous solution was $-514\times10^{-10}$ for the Liebmann procedure and only $700\times10^{-18}$ for the seventh-order procedure. When the net constant was doubled the maximum error for the Liebmann procedure was $-1940\times10^{-10}$, approximately four times the previous value, in accordance with (2). For the seventh-order procedure the error was $46{,}000\times10^{-18}$ with $h=2$, approximately 64 times the previous error, since with this more accurate interpolation formula the error varies as $h^6$ in place of $h^2$. This is another great advantage of this procedure in cases where it can be used.

On this same example, with $h=1$, the maximum error for the nine-block procedure was $+108\times10^{-10}$, and in each block there were both positive and negative errors while the Liebmann errors were everywhere larger and negative. These facts are explained by the fact that for $\log r^2$ in this region the value at a point is always greater than the average of its four neighbors, but less than the average of its four diagonal neighbors. The seventh-order formula (24) corrects for this by the proper mixture of averages taken in the two ways. The nine-block procedure apparently tends to make this correction on the average by calculating five of the points in the Liebmann way but the corner points diagonally. A similar type of correction of the error in the Liebmann procedure is possible in all cases, and it may be that in general the nine-block procedure will give a closer approximation than the Liebmann because of this partial correction.

### d. Proof of convergence of the iteration process

Each of the single-point improvement formulas one uses is of the type (21) in which the improved value at a point is a weighted average ($\sum b_i = 1$) of the values, some already improved, some unimproved, of the function at neighboring points. This is seen to imply that the improved value $\bar{\psi}_i$ of (4) is in the last analysis a true weighted *average* of the values of $\psi$ at all interior and boundary points, i.e. that

$$\sum_j T_{ij} + \sum_\alpha U_{i\alpha} = 1.$$

The considerations of §2 furnish a proof of the convergence of the iteration process provided that we can prove that the magnitude of the greatest eigenvalue of $\boldsymbol{T}$ is less than unity. This will be proved if we can show that $\boldsymbol{T}$ applied to any function whatsoever of zero boundary values reduces the absolute maximum of the function. This is easily shown provided that all of the $b$'s of the improve-

ment formula (21) are positive, since then all of the $T_{ij}$'s and $U_{i\alpha}$'s of (4) will be positive. Although it is conceivable that cases might arise in which negative values of $b$ appear desirable, this is not the case in any situation which we have discussed or contemplated, and so seems unlikely. With positive $b$'s, the maximum value of a function is replaced upon improvement by a value which is a weighted mean (with all weights positive) of the values of the function at all interior points and of the zero values of the function at the boundary points and this clearly reduces this maximum value.[15] No other point can get replaced by a value so large as this maximum, so the maximum of any function is definitely reduced upon operation with $\boldsymbol{T}$. This is true in particular of the highest eigenfunction, so the highest eigenvalue must be less than unity.

### e. Calculation of asymptotic eigenvalues

We shall illustrate the method of calculating the asymptotic form of the eigenvalues as the number of net points increases by treating the case first discussed in the text in which the improvement is by blocks of $v^2$ points using such formulae that the difference equation (1) is satisfied at each point of the block by the improved values. [The eigenvalues for the more practical block procedures can be obtained in much the same way; we shall indicate in brackets the modifications of the calculation which are necessary in these cases.]

We shall consider first a square region containing $m$ blocks on a side. Letter the points according to their location in the block, and number the blocks with two indices according to their position in the net, as in Fig. 7 which applies to $v=3$. We denote the values of the eigenfunction $\phi^{(1)}$ at the points $a_{ij}$, $b_{ij}$, $\cdots$, simply by $a_{ij}$, $b_{ij}$; $\cdots$; the values of the improved function $\bar{\phi}^{(1)}$ by $\bar{a}_{ij}$, $\bar{b}_{ij}$, $\cdots$. If the blocks are improved in the order 11, 21, 31, $\cdots$; 12, 22, 32, $\cdots$; etc., the improved values in the $ij$ block satisfy equations similar to the following, which are for $v=3$:

$$\begin{aligned}
-4\bar{a}_{ij}+\bar{f}_{ij}+\bar{g}_{ij}+\bar{h}_{ij}+\bar{l}_{ij}&=0,\\
-4\bar{b}_{ij}+\bar{f}_{ij}+\bar{l}_{ij}+\bar{c}_{i-1,\,j}+\bar{e}_{i,\,j-1}&=0,\\
-4\bar{c}_{ij}+\bar{g}_{ij}+\bar{f}_{ij}+\bar{d}_{i,\,j-1}+b_{i+1,\,j}&=0,\\
-4\bar{d}_{ij}+\bar{h}_{ij}+\bar{g}_{ij}+e_{i+1,\,j}+c_{i,\,j+1}&=0,\\
-4\bar{e}_{ij}+\bar{l}_{ij}+\bar{h}_{ij}+b_{i,\,j+1}+\bar{d}_{i-1,\,j}&=0,\\
-4\bar{f}_{ij}+\bar{a}_{ij}+\bar{b}_{ij}+\bar{c}_{ij}+\bar{h}_{i,\,j-1}&=0,\\
-4\bar{g}_{ij}+\bar{a}_{ij}+\bar{c}_{ij}+\bar{d}_{ij}+l_{i+1,\,j}&=0,\\
-4\bar{h}_{ij}+\bar{a}_{ij}+\bar{d}_{ij}+\bar{e}_{ij}+f_{i,\,j+1}&=0,\\
-4\bar{l}_{ij}+\bar{a}_{ij}+\bar{e}_{ij}+\bar{b}_{ij}+\bar{g}_{i-1,\,j}&=0.
\end{aligned}\qquad(25)$$

[For other procedures one writes here the actual equations used in the improvement process, which are usually more complicated.]

The condition that $\phi$ represent an eigenfunction is that

$$\bar{a}_{ij}=\lambda a_{ij},\quad \bar{b}_{ij}=\lambda b_{ij},\quad \text{etc.}\qquad(26)$$

[15] For a function in which the maximum value is assumed at several neighboring points, it may be that for some order of improvement, *not all* of these maximum points will be reduced the first time over. They will all be reduced after several times over. The eigenfunctions are obviously not of this type, so this does not affect the convergence proof.

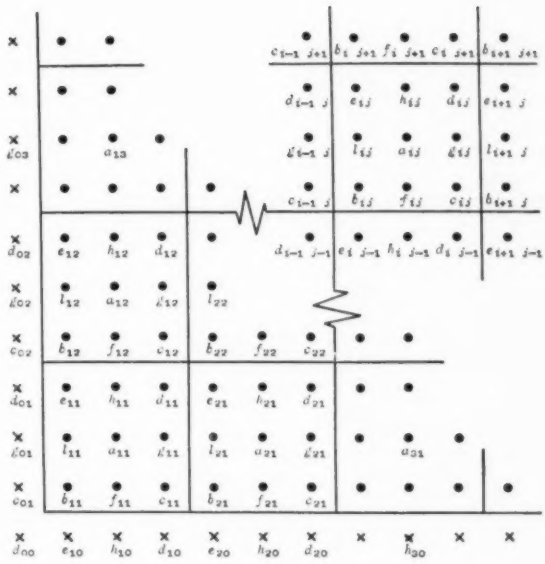


FIG. 7.

We shall not solve directly the problem of $m^2$ blocks with a row of zero values around the outside since these boundary conditions are difficult to express analytically. Rather we shall assume that the function value at a particular type of point, say at the $b$ points, is of the form

$$b_{ij}=\left(1+B\frac{\pi^2}{m^2}\right)\sin\frac{\pi}{m}(i+x_b)\sin\frac{\pi}{m}(j+y_b),\qquad(27)$$

to terms of order $1/m^2$; here $B$, $x_b$, and $y_b$ are constants of the order of unity. Assuming a function of this form for each type of point amounts to replacing our problem by a periodic problem with "boundary" conditions $a_{i,\,j}=-a_{i,\,j+m}=-a_{i+m,\,j}$; $b_{i,\,j}=-b_{i,\,j+m}=-b_{i+m,\,j}$; etc. The choice (27) amounts to placing the "nodes" of the function in the blocks with $i$ or $j=0$, $m$, $2m$, $\cdots$; in these blocks the function is of the order of $1/m$ times its maximum value. For these homogeneous boundary conditions, $(1-\lambda_1)m^2$ will have the same asymptotic value as for the zero boundary conditions used in practice.

Since $\lambda_1=1-O(1/m^2)=1-k\pi^2/m^2+O(1/m^3)$, (28)

we need only satisfy the equations (25) to terms of the order of $1/m^2$ in order to find the asymptotic value $k\pi^2$ of $(1-\lambda_1)m^2$. It turns out that it is possible to do this with functions of the type (27). We expand (27) in the form

$$\begin{aligned}
b_{i+\bullet,\,j+\star}=&\{1+[B-\tfrac{1}{2}(x_b+\bullet)^2-\tfrac{1}{2}(y_b+\star)^2]\pi^2/m^2\}\\
&\times\sin i\pi/m\sin j\pi/m\\
&+\{(x_b+\bullet)\pi/m\}\cos i\pi/m\sin j\pi/m\\
&+\{(y_b+\star)\pi/m\}\sin i\pi/m\cos j\pi/m\\
&+\{(x_b+\bullet)(y_b+\star)\pi^2/m^2\}\cos i\pi/m\cos j\pi/m+O(1/m^3),
\end{aligned}\qquad(29)$$

with similar forms for $a_{i+\bullet,\,j+\star}$, $c_{i+\bullet,\,j+\star}$, etc. Here $\bullet=+1$, 0, or $-1$; $\star=+1$, 0, or $-1$.

These expressions are to be substituted in the equations such as (25), and the coefficients of the sin sin, sin cos, cos sin, and cos cos terms to be equated to zero. From the form of (25) and the coefficients of cos sin in (29) we see

that the $x+\bullet$ values must increase in uniform steps as we move to the right, so that, except for an additive constant,

$$x_d-1=x_g-1=x_c-1=-\tfrac{2}{3},$$
$$x_e=x_i=x_b=-\tfrac{1}{3},$$
$$x_h=x_a=x_f=0,$$
$$x_d=x_g=x_e=+\tfrac{1}{3},$$
$$x_e+1=x_i+1=x_b+1=+\tfrac{2}{3}$$

for $v=3$. In general the corresponding values will be $(-v-1)/2v$, $(-v+1)/2v$, $\cdots$, $(v-1)/2v$, $(v+1)/2v$, arranged similarly in the blocks. The $y$ values are the same as the $x$ values with the block turned through 90°. With these values it is easily verified that the cos cos equations are automatically satisfied. [For all other procedures for which we have calculated asymptotic values, these same $x$ and $y$ values have satisfied the cos sin, sin cos, and cos cos equations.]

Finally, equating the coefficients of the sin sin terms furnishes a set of $v^2$ linear equations to determine the $v^2$ constants $A$, $B$, $C$, $\cdots$, and also $k$. We see that each of the capital letters occurs $-4$ times in one equation and once in four equations, so that if the whole $v^2$ equations be added $A$, $B$, $C$, $\cdots$ drop out leaving a single equation for $k$. [The equations obtained from setting the coefficients of sin sin equal to zero are in general unchanged by adding the same constant to $A$, $B$, $C$, $\cdots$—this corresponds to multiplication of $\phi$ by a constant of the magnitude of $1+O(1/m^2)$—so we can assume one arbitrary relation between them. In the twenty-five block procedure it was necessary to assume the sum of the values at the four $a$ points of Fig. 3 to be zero before a linear combination of the equations could be found which eliminated all unknowns but $k$. In the other cases we have considered, a satisfactory linear combination could be made without such an assumption.] From (26) we see that $-k\pi^2/m^2$ occurs in this equation once for each barred value so $k$ occurs once for each unbarred value. Since unbarred values occur only on two edges of the block, the coefficient of $k\pi^2/m^2$ in the resultant equation is $2v$. Finally we must calculate the constant term. Since each *letter* occurs $-4$ times and $+4$ times in (25), the $x^2$ and $y^2$ parts of (29) cancel, leaving only the terms involving $\bullet$ and $\bigstar$. The $\bullet$ terms involve $+x$ for each of the $v$ points bordering the $ij$ block to the left and $-x$ for the points bordering to the right; giving in particular for $v=3$, $x_d+x_g+x_c-x_e-x_i-x_b=6\cdot\frac{1}{3}=2$; in general, $2v\cdot(v-1)/2v=(v-1)$ in units $\pi^2/m^2$. One has also $-\frac{1}{2}\bullet^2$ for each of these points, or $-v$ altogether. The whole contribution of the $\bullet$ terms is then $-1$; the $\bigstar$ terms also give $-1$, total $-2\pi^2/m^2$.

The equation resulting upon addition of all the equations of the type (25) is then

$$2vk-2=0$$

or
$$1-\lambda_1=k\pi^2/m^2=\pi^2/vm^2=v\pi^2/p^2,$$

where $p=vm$ (asymptotically) is the number of intervals on a side. This is the result discussed in the text.

For a rectangular region one obtains a value equal to (14b) with $v\pi^2$ written for $\pi^2$. One can derive this by the method just outlined if one takes $m$ blocks in one direction, $l$ blocks in the other. Asymptotic values for the other eigenfunctions can be shown to be given by (14c) with $v\pi^2$ written for $\pi^2$ by taking in place of (27) a product of $\sin(\mu\pi/m)(i+x_b)$ $\sin(\nu\pi/l)(j+y_b)$. [For any other procedure, (14a), (14b) and (14c) are also obtained, with const. $\times\pi^2$ replacing $\pi^2$.]

Finally we might remark that if in the improvement of a block the original $\phi$ values are used for all boundary values, in place of the improved value $\bar{\phi}$ where they are available; e.g. if in (25) the bars are removed from letters with subscripts $i-1$, $j$ and $i$, $j-1$; the asymptotic values of $1-\lambda$ come out only half as great as before. This is true of any procedure in which the points of a block are treated symmetrically. Hence we see that the device of using improved values as fast as they are available in the improvement of other points just *doubles* the rate of convergence over a similar procedure in which improved values are not used.